REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
02-02-2016	Final Report	1-Aug-2009 - 31-May-2015

4. TITLE AND SUBTITLE
Final Report: Unified Research on Network-Based Hard/Soft Information Fusion

5a. CONTRACT NUMBER
W911NF-09-1-0392

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611103

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS
Dr. Rakesh Nagi, Dr. David Hall (late), Dr. Amir Shirkhodaie, Dr. Ronald Yager, Dr. Moises Sudit, Dr. James Llinas, Dr. Ann Bisantz, Dr. Stuart Shapiro, Dr. Geoff Gross, Mr. Ketan Date

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
State University of New York (SUNY) at Bu
Sponsored Projects Services
402 Crofts Hall
Buffalo, NY 14260 -7016

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
56191-MA-MUR.145

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The University at Buffalo (UB) Center for Multisource Information Fusion (CMIF) and its partners including the Pennsylvania State University (PSU), Iona College (Iona), Tennessee State University (TSU) and University of Illinois at Urbana-Champaign (UIUC) have conducted research to develop a generalized framework, mathematical techniques, and test and evaluation methods to address the ingestion and harmonized fusion of Hard and Soft information in a distributed (networked) Level 1 and Level 2 data fusion environment. This research activity is supported by a Multidisciplinary University Research Initiative

15. SUBJECT TERMS
Information Fusion, Hard-Soft Information Fusion, Graph Association, Data Association, Natural Language Understanding, Test and Evaluation, Counter-Insurgency, Big Data and Big Graphs

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UU	UU	UU	UU		Rakesh Nagi
					19b. TELEPHONE NUMBER
					716-645-2427

Report Title

Final Report: Unified Research on Network-Based Hard/Soft Information Fusion

ABSTRACT

The University at Buffalo (UB) Center for Multisource Information Fusion (CMIF) and its partners including the Pennsylvania State University (PSU), Iona College (Iona), Tennessee State University (TSU) and University of Illinois at Urbana-Champaign (UIUC) have conducted research to develop a generalized framework, mathematical techniques, and test and evaluation methods to address the ingestion and harmonized fusion of Hard and Soft information in a distributed (networked) Level 1 and Level 2 data fusion environment. This research activity is supported by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for "Unified Research on Network-based Hard/Soft Information Fusion," issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery and recently Dr. Joe Myers. This report provides a summary of the five years of progress. The primary Research Thrusts addressed are framed around the major functional components of the JDL Fusion Process; these include:

1. Source Characterization of Soft Data input streams including; human observation.—
direct, indirect, open source inputs, linguistic framing, and text processing
2. Common Referencing and Alignment of Hard and Soft Data, especially strategies and
methods for meta-data generation for Hard-Soft data normalization
3. Generalized Data Association Strategies and Algorithms for Hard and Soft Data
4. Robust Estimation Methods that exploit associated Hard and Soft Data
5. Dynamic Network-based Effects on Hard-Soft Data Fusion Architectures and Methods
6. Test and Evaluation Methodology Development to include Human-in-the-Loop
7. Extensibility, Adaptability, and Robustness Assessment
8. Fusion Process Framework
9. Technology Concept of Employment

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Received	Paper
02/02/2016 44.00	Rakesh Nagi, Geoff A. Gross. Precedence tree guided search for the efficient identification of multiple situations of interest – AND/OR graph matching, Information Fusion, (01 2016): 240. doi: 10.1016/j.inffus.2015.02.001
08/29/2013 77.00	Geoff A. Gross, Ann M. Bisantz, Rakesh Nagi, Michael P. Jenkins. Towards context aware data fusion: Modeling and integration of situationally qualified human observations to manage uncertainty in a hard+soft fusion process, Information Fusion, (06 2013): 0. doi: 10.1016/j.inffus.2013.04.011
08/29/2013 85.00	Ronald R. Yager. On Prioritized Multiple-Criteria Aggregation, , (10 2012): 0. doi:
08/29/2013 86.00	Ronald Yager. Participatory Learning of Propositional Knowledge, , (08 2012): 0. doi:
08/29/2014 03.00	Geoff Gross, Rakesh Nagi, Kedar Sambhoos. A fuzzy graph matching approach in intelligence analysis and maintenance of continuous situational awareness, Information Fusion, (07 2014): 0. doi: 10.1016/j.inffus.2013.05.006
08/29/2014 01.00	Gregory Tauer, Rakesh Nagi, Moises Sudit. The graph association problem: Mathematical models and a lagrangian heuristic, Naval Research Logistics (NRL), (04 2013): 0. doi: 10.1002/nav.21532
08/29/2014 02.00	Gregory Tauer, Rakesh Nagi. A map-reduce lagrangian heuristic for multidimensional assignment problems with decomposable costs, Parallel Computing, (11 2013): 0. doi: 10.1016/j.parco.2013.08.012
08/30/2014 15.00	Ronald R. Yager. Pythagorean Membership Grades in Multicriteria Decision Making, IEEE Transactions on Fuzzy Systems, (08 2014): 0. doi: 10.1109/TFUZZ.2013.2278989
08/30/2014 16.00	Ronald R. Yager, Naif Alajlan. Probability weighted means as surrogates for stochastic dominance in decision making, Knowledge-Based Systems, (08 2014): 0. doi: 10.1016/j.knosys.2014.04.024
08/30/2014 17.00	Ronald R. Yager, Naif Alajlan. Probabilistically Weighted OWA Aggregation, IEEE Transactions on Fuzzy Systems, (02 2014): 0. doi: 10.1109/TFUZZ.2013.2245899
08/30/2014 18.00	Ronald R. Yager, Naif Alajlan. On characterizing features of OWA aggregation operators, Fuzzy Optimization and Decision Making, (07 2013): 0. doi: 10.1007/s10700-013-9167-8
08/30/2014 19.00	Ronald R. Yager. Joint cumulative distribution functions for Dempster–Shafer belief structures using copulas, Fuzzy Optimization and Decision Making, (05 2013): 0. doi: 10.1007/s10700-013-9163-z
08/30/2014 20.00	Ronald R. Yager, Naif Alajlan. Decision Making with Ordinal Payoffs Under Dempster-Shafer Type Uncertainty, International Journal of Intelligent Systems, (11 2013): 0. doi: 10.1002/int.21615

08/30/2014 22.00 Ronald R. Yager, Naif Alajlan. A note on mean absolute deviation, Information Sciences, (09 2014): 0. doi: 10.1016/j.ins.2014.04.016

TOTAL: **14**

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper
08/29/2013 87.00	Ronald Yager. On a View of Zadeh's Z-Numbers, Advances in Computational Intelligence, (08 2012): 0. doi:
08/29/2013 88.00	Ronald R. Yager, Naif Alajlan. Measure based representation of uncertain information, Fuzzy Optimization Decision Making, (08 2012): 0. doi:
08/29/2013 89.00	Ronald Yager. Membership modification and level sets, Soft Computing A Fusion of Foundations,Methodologies and Applications, (09 2012): 0. doi:
08/29/2013 90.00	Ronald Yager. On forming joint variables in computing with words, International Journal of General Systems, (07 2013): 0. doi:
08/29/2013 91.00	Ronald R. Yager, Ali M. Abbasov. Pythagorean Membership Grades,Complex Numbers, and Decision Making, International Journal of Intelligent Systems, (05 2013): 0. doi:

TOTAL: **5**

Number of Papers published in non peer-reviewed journals:

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received		Paper
08/28/2012	41.00	Amir Shirkhodaie, Vinayak Elagovan, Aaron Rababaah. Acoustic Semantic Labeling and Fusion of Human-Vehicle Interactions, 2nd Annual Human & Light Vehicle & Tunnel Detection Workshop. 03-MAY-11, . : ,
08/28/2012	44.00	James Llinas. Situation Management in Counter-Insurgency Operations: An Overview of Operational Art and Relevant Technologies, 2011 Proceedings of the 14th International Conference on Information Fusion. 04-JUL-11, . : ,
08/28/2012	42.00	Vinayak Elangovan, Amir Shirkhodaie. Adaptive Characterization, Tracking, and Semantic Labeling of Human-Vehicle Interactions Via Multi-modality Data Fusion, SPIE Electronic Imaging Science and Technology Conference, Multimedia Content Access: Algorithms and Systems. 22-JAN-12, . : ,
08/31/2011	20.00	Jacob L. Graham, , David L. Hall, , Jeffrey Rimland. A new synthetic dataset for evaluating soft and hard fusion algorithms, Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011. , . : ,
10/11/2011	26.00	D. J. Natale, M. S. Baran, R. L. Tutwiler, R. Tutwiler. Live Motion 3D Data Processing, IDGA 9th Image Fusion Summit. , . : ,
10/11/2011	32.00	Roanld Yager. Criteria satisfaction under measure based uncertainty, Fuzzy Optimization and Decision Making. , . : ,

TOTAL: **6**

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received Paper

08/28/2012 37.00 Ronald Yager. Conditional approach to possibility-probability fusion,
IEEE TRANSACTIONS ON FUZZY SYSTEMS. 01-FEB-12, . : ,

08/28/2012 51.00 Daniel R. Schlegel, Stuart C. Shapiro. Visually Interacting with a Knowledge BaseUsing Frames, Logic, and Propositional Graphs,
Graph Structures for Knowledge Representation and ReasoningLecture Notes in Computer Science. 01-AUG-11, . : ,

08/28/2012 50.00 Geoff A. Gross, Rakesh Nagi, Kedar Sambhoos, Daniel R. Schlegel, Stuart C. Shapiro, Gregory Tauer. Towards Hard+Soft Data Fusion: ProcessingArchitecture and Implementation for the Joint Fusionand Analysis of Hard and Soft Intelligence Data,
15th International Conference on Information Fusion. 09-JUL-12, . : ,

08/28/2012 48.00 Vinayak Elangovan, Vinod K Bandaru, Amir Shirkhodaie. Team Activity Analysis and Recognition Based on Kinect Depth Map and Optical Imagery Techniques,
SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI. 23-APR-12, . : ,

08/28/2012 47.00 Amjad Alkilani, Amir Shirkhodaie. A Survey on Acoustic Signature Recognition and Classification Techniques For Persistent Surveillance Systems,
SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI. 23-APR-12, . : ,

08/28/2012 46.00 Mohammad S. Habibi, Amir Shirkhodaie. A Survey of Visual Analytics for Knowledge Discovery andContent Analysis,
SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI. 23-APR-12, . : ,

08/28/2012 45.00 Vinayak Elangovan, Amir Shirkhodaie. Recognition of Human Activity Characteristics Based onState Transitions Modeling Technique,
SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI. 23-APR-12, . : ,

08/28/2012 43.00 James Llinas, Jeffrey C. Rimland. Network and Infrastructure Considerations for Hardand Soft Information Fusion Processes,
5th International Conference on Information Fusion. 09-JUL-12, . : ,

08/29/2012 53.00 Sushant S. Khopkar, Rakesh Nagi, Alexander G. Nikolaev. An Efficient Map-Reduce Algorithm for theIncremental Computation of All-Pairs ShortestPaths in Social Networks,
First International Workshopon Data Management in the Social Semantic Web. 26-AUG-12, . : ,

08/29/2012 60.00 Ronald R. Yager. A measure based approach to the fusion of possibilisticand probabilistic uncertainty,
Fuzzy Optim Decis Making. 13-MAR-11, . : ,

08/29/2012 58.00 Jeffrey C. Rimland, David L. Hall, Jacob L. Graham. Human cognitive and perceptual factors in JDL level 4 hard / soft data fusion,
Proceedings of SPIE. 23-APR-12, . : ,

08/29/2012 57.00 Jeffrey C. Rimland, Dan Coughlin, David L. Hall, Jacob L. Graham. Advances in data representation for hard/soft information fusion,
Proceedings of SPIE. 23-APR-12, . : ,

08/29/2012 54.00 Matthew S. Baran, Donald J. Natale, Richard L. Tutwiler, Christopher Griffin, John M. Daughtry, Matthew M. McQuillan, Jeffrey Rimland, David L. Hall. Hard Sensor Fusion for COIN Inspired SituationAwareness,
2011 Proceedings of the 14th International Conference on Information Fusion (FUSION). 05-JUL-11, . : ,

08/29/2013 78.00 Michael Stearns, Alexander Nikolaev, Sue Kase, Kirk Ogaard. Fusion of a Set of Attributed Graphs for Event Reconstruction,
Proceedings of IIE ISERC 2013 Conference. 01-MAY-13, . : ,

08/29/2013 76.00 Sushant Khopkar, Rakesh Nagi, Kedar Sambhoos, Ketan Date, Geoff A. Gross. Data Association and Graph Analytical Processing ofHard and Soft Intelligence Data,
International Conference on Information Fusion. 07-JUL-13, . : ,

08/29/2013 72.00 Jeffrey C. Rimland, Dan Coughlin, David L. Halla, Jacob L. Grahama. Advances in data representation for hard/soft information fusion,
Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2012. 23-APR-12, . : ,

08/29/2013 70.00 Jeffrey Rimlanda, Mark Ballorab, Wade Shumakera. Beyond visualization of big data: a multi-stage data exploration approach using visualization, sonification, and storification,
SPIE Defense, Security, and Sensing Symposium. 01-OCT-13, . : ,

08/29/2013 69.00 Jeffrey Rimland, James Llinas. Network and Infrastructure Considerations for Hard and Soft Information Fusion Processes,
International conference on Information Fusion. 07-JUL-12, . : ,

08/29/2013 74.00 Jeffrey Rimland, Michael McNeese, David Hall. Conserving analyst attention units: use of multi-agent software and CEP methods to assist information analysis,
Proc. SPIE 8758, Next-Generation Analyst,. 28-MAY-13, . : ,

08/29/2013 73.00 Jeffrey C. Rimland, David L. Hall, Jacob L. Graham. Human cognitive and perceptual factors in JDL level 4 hard / soft data fusion,
Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2012. 25-JUN-12, . : ,

08/29/2013 79.00 James Llinas. Challenges in Information Fusion Technology Capabilities for Modern Intelligence and Security Problems,
European Intelligenceand Security Informatics Conference (EISIC) 2013. 12-AUG-12, . : ,

08/29/2013 81.00 Daniel R. Schlegel, Stuart C. Shapiro. Concurrent Reasoning withInference Graphs,
Proceedings of the Twenty-Seventh AAAI Conference onArtificial Intelligence (AAAI-13). 01-MAY-13, . : ,

08/29/2013 80.00 Stuart C. Shapiro , Daniel R. Schlegel. Natural Language Understandingfor Soft Information Fusion,
Proceedings ofthe 16th International Conference on Information Fusion (Fusion 2013). 07-JUL-13, . : ,

08/29/2013 82.00 Daniel Schegel. Concurrent Inference Graphs,
Proceedings of the Twenty-Seventh AAAIConference on Artificial Intelligence (AAAI-13). 01-MAY-13, . : ,

08/29/2013 84.00 Michael P. Jenkins, Ann Bisantz, James Llinas, Rakesh Nagi. Investigating and Improving Network Visualizations'Effectiveness at Supporting Human Sensemaking Tasks,
57th Annual Meeting of the Human Factors and Ergonomics Society (HFES). 01-MAY-13, . : ,

08/29/2013 92.00 Ronald Yager, Fred Petry. Intuitive Decision-Making Using Hyper SimilarityMatching,
2013 IFSA World CongressNAFIPS Annual Meeting. 24-JUN-13, . : ,

08/29/2013 93.00 Nathan S. Butler, Richard L. Tutwiler . Invariant Unsupervised Segmentation ofDismounts in Depth Images,
Signal Processing, Sensor Fusion, and Target Recognition. 23-MAY-13, . : ,

08/29/2013 94.00 Matthew S. Baran, Richard L. Tutwiler, Donald J. Natale, Michael S. Bassett, Matthew P. Harner, Sylvia S. Shen, Paul E. Lewis. Multimodal detection of man-made objects in simulated aerial images,
SPIE Defense, Security, and Sensing. 18-MAY-13, Baltimore, Maryland, USA. : ,

08/29/2014 04.00 James Llinas. Reexamining information fusion-decision making inter-dependencies,
2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). 02-MAR-14, San Antonio, TX, USA. : ,

08/29/2014 10.00 Dana Kerker, Michael P. Jenkins, Geoff A. Gross, Ann M. Bisantz, Rakesh Nagi. Visual estimation of human attributes: An empirical study of context-dependent human observation capabilities,
2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). 02-MAR-14, San Antonio, TX, USA. : ,

08/29/2014 09.00 Daniel R. Schlegel, Stuart C. Shapiro. The ‘Ah Ha!’ Moment : When Possible, Answering the Currently Unanswerableusing Focused Reasoning,
36th Annual Conference of the Cognitive Science Society. 23-JUL-14, . : ,

08/29/2014 08.00 Daniel R. Schlegel, Stuart C. Shapiro. Inference Graphs: A New Kind of Hybrid Reasoning System,
Twenty-Eighth AAAI Conference on Artificial Intelligence . 27-JUL-14, . : ,

08/29/2014 07.00 Geoff A. Gross, Ketan Date, Daniel R. Schlegel, Jason J. Corso, James Llinas, Rakesh Nagi, Stuart C. Shapiro. Systemic Test and Evaluation of a Hard+Soft Information Fusion Framework,
17th International Conference on Information Fusion. 07-JUL-14, . : ,

08/29/2014 06.00 Ketan Date, Geoff A. Gross, Rakesh Nagi. Test and Evaluation of Data Association Algorithms in Hard+Soft Data Fusion,
17th International Conference on Information Fusion. 07-JUL-14, . : ,

08/29/2014 05.00 James Llinas. Challenges in Information Fusion Technology Capabilities for Modern Intelligence and Security Problems,
2013 European Intelligence and Security Informatics Conference (EISIC). 11-AUG-13, Uppsala, Sweden. : ,

08/30/2013 98.00 Jake Graham, David Hall. The use of Analytic Decision Game (ADG) methods for test and evaluation of hard and soft data fusion systems and education of a new generation of data fusion analysts,
Proceedings of the National Symposium on Sensor Data Fusion (NSSDF). 22-OCT-12, . : ,

08/30/2013 00.00 Donald R. Kretz, B. J. Simpson, Colonel Jacob Graham. A game-based experimental protocol for identifying and overcoming judgment biases in forensic decision analysis,
2012 IEEE International Conference on Technologies for Homeland Security (HST). 13-NOV-12, Waltham, MA, USA. : ,

08/30/2014 13.00 Daniel R. Schlegel, Stuart C. Shapiro. Inference Graphs: A Roadmap,
Second Annual Conference on Advances in Cognitive Systems. 12-DEC-13, . : ,

08/30/2014 12.00 James Llinas. A survey of automated methods for sensemaking support,
SPIE Sensing Technology + Applications. 05-MAY-14, Baltimore, Maryland, USA. : ,

08/30/2014 14.00 Guoray Cai, Geoff Gross, James Llinas, David Hall. A visual analytic framework for data fusion in investigative intelligence,
SPIE Sensing Technology + Applications. 05-MAY-14, Baltimore, Maryland, USA. : ,

08/30/2014 24.00 Ronald R. Yager. Pythagorean fuzzy subsets,
2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS). 23-JUN-13, Edmonton, AB, Canada. : ,

08/30/2014 25.00 Rachel L. Yager, Ronald R. Yager. Soft Retrieval and Uncertain Databases,
2014 47th Hawaii International Conference on System Sciences (HICSS). 05-JAN-14, Waikoloa, HI, USA. : ,

08/30/2014 29.00 Jeff Rimland, David Hall. Hard and soft information fusion in sonification for assistive mobile device technology,
International Conference on Auditory Display. 06-JUL-13, . : ,

08/30/2014 30.00 Matthew S. Baran, Richard L. Tutwiler, Donald J. Natale, Michael S. Bassett, Matthew P. Harner, Sylvia S. Shen, Paul E. Lewis. Multimodal detection of man-made objects in simulated aerial images,
SPIE Defense, Security, and Sensing. 18-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 31.00 Richard L. Tutwiler, Ivan Kadar, Nathan S. Butler. Invariant unsupervised segmentation of dismounts in depth images,
SPIE Defense, Security, and Sensing. 23-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 32.00 Jeffrey C. Rimland, Michael McNeese, David L. Hall. Conserving analyst attention units: use of multi-agent software and CEP methods to assist information analysis,
SPIE Defense, Security, and Sensing. 28-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 33.00 Steven C. Shaffer. Automatic theory generation from analyst text files using coherence networks,
SPIE Sensing Technology + Applications. 22-MAY-14, Baltimore, Maryland, USA. : ,

08/30/2014 34.00 John P. Morgan, Richard L. Tutwiler. Real-time reconstruction of depth sequences using signed distance functions,
SPIE Defense + Security. 20-JUN-14, Baltimore, Maryland, USA. : ,

08/30/2014 35.00 Amjad Alkilani, Amir Shirkhodaie. Acoustic events semantic detection, classification, and annotation for persistent surveillance applications,
SPIE Defense + Security. 20-JUN-14, Baltimore, Maryland, USA. : ,

08/30/2014 36.00 Vinayak Elangovan, Amir Shirkhodaie. Knowledge discovery in group activities through sequential observation analysis,
SPIE Defense + Security. 20-JUN-14, Baltimore, Maryland, USA. : ,

08/30/2014 37.00 Mohammad S. Habibi, Amir Shirkhodaie. Multi-attributed tagged big data exploitation for hidden concepts discovery,
SPIE Defense + Security. 20-JUN-14, Baltimore, Maryland, USA. : ,

08/30/2014 38.00 Amir Shirkhodaie, Vinayak Elangovan, Amjad Alkilani, Mohammad Habibi. A decision support system for fusion of hard and soft sensor information based on probabilistic latent semantic analysis technique,
SPIE Defense, Security, and Sensing. 28-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 39.00 Vinayak Elangovan, Amir Shirkhodaie. A robust technique for semantic annotation of group activities based on recognition of extracted features in video streams,
SPIE Defense, Security, and Sensing. 23-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 40.00 Vinayak Elangovan, Bashir Alsaidi, Amir Shirkhodaie. A multi-attribute based methodology for vehicle detection and identification,
SPIE Defense, Security, and Sensing. 23-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 41.00 Amjad Alkilani, Amir Shirkhodaie. Acoustic signature recognition technique for Human-Object Interactions (HOI) in persistent surveillance systems,
SPIE Defense, Security, and Sensing. 23-MAY-13, Baltimore, Maryland, USA. : ,

08/30/2014 42.00 Mohammad S. Habibi, Amir Shirkhodaie. Mining patterns in persistent surveillance systems with smart query and visual analytics,
SPIE Defense, Security, and Sensing. 23-MAY-13, Baltimore, Maryland, USA. : ,

08/31/2011 2.00 Jenkins, M.P., Bisantz, A.M. Identification of Human-Interaction Touch Points for Intelligence Analysis Information Fusion Systems,
14th International Conference on Information Fusion, Special Session on Hard/Soft Information Fusion: New Data Sets and Innovative Architectures. 05-JUL-11, . : ,

08/31/2011 40.00 James Llinas . Situation Management in Counter-Insurgency Operations: An Overview of Operational Art and Relevant Technologiess,
14th International Conference on Information Fusion. , . : ,

08/31/2011 39.00 Deven McMaster, Rakesh Nagi, Kedar Sambhoos. Temporal Alignment in Soft Information Process,
14th International Conference on Information Fusion. , . : ,

08/31/2011 38.00 Megan Hannigan, , James Llinas,, Kedar Sambhoos .
Specificity and Merging Challenges in Soft Data Association ,
14th International Conference on Information Fusion. , . : ,

08/31/2011 27.00 Geoff Gross,, Rakesh Nagi, Kedar Sambhoos. Continuous Preservation of Situational Awareness through Incremental and Stochastic Graphical Methods,
14th International Conference on Information Fusion. Proc. of Fusion 2011. , . : ,

08/31/2011 25.00 J. Graham, J. Rimland, D. Hall. A COIN-inspired synthetic data set for quantitative evaluation of hard and soft fusion systems,
Proceedings of Fusion 2011: the International Conference on Information Fusion. , . : ,

08/31/2011 24.00 R. Tutwiler, M. Baran, D. Natale, C. Griffin, J. Daughtry, M. McQuillan, J. Rimland, D. Hall. Hard sensor fusion for COIN inspired situation awareness,
Proceedings of Fusion 2011: the International Conference on Information Fusion. , . : ,

08/31/2011 23.00 Jeffrey C Rimland, David L. Hall. A multi-agent infrastructure for hard and soft information fusion,
Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011. , . : ,

08/31/2011 22.00 Jeffrey C Rimland, Ganesh M Iyer, Rachana R Agumamidi, Soumya V Pisupati, Jake Graham. JDL level 0 and 1 algorithms for processing and fusion of hard sensor data,
Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011. , . : ,

08/31/2011 21.00 D. J. Natale, M. S. Baran, R. Tutwiler , D. L. Hall. 3DSF: three dimensional spatio-temporal fusion,
Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011. , . : ,

08/31/2011 3.00 Gross, G., Bisantz, A.M., Nagi, R, Jenkins, M.P.. Towards context-aware hard/soft information fusion: Incorporating situationally qualified human observations into a fusion process for intelligence analysis,
2011 IEEE First International Multi-Disciplinary Conference on . 22-FEB-11, . : ,

08/31/2011 5.00 Stuart C. Shapiro. The Jobs Puzzle: A Challenge for Logical Expressibility and Automated Reasoning,
Logical Formalizations of Commonsense Reasoning: Papers from the AAAI Spring Symposium. , . : ,

08/31/2011 6.00 Michael Kandefer, Stuart C. Shapiro. Evaluating Spreading Activation for Soft Information Fusion,
14th International Conference on Information Fusion. , . : ,

08/31/2011 7.00 Stuart C. Shapiro, Michael Prentice . Using Propositional Graphs for Soft Information Fusion,
14th International Conference on Information Fusion . 05-JUL-11, . : ,

08/31/2011 8.00 Stuart C. Shapiro, Daniel R. Schlegel. Visually Interacting with a Knowledge Base Using Frames, Logic, and Propositional Graphs,
Second International IJCAI Workshop on Graph Structures for Knowledge Representation and Reasoning. , . : ,

10/11/2011 9.00 Elangovan V., Shirkhodaie, A., . A Survey of Imagery Techniques for Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,
SPIE Defense and Security Conference. , . : ,

10/11/2011 10.00 Shirkhodaie, A., Rababaah, A., Elangovan V.,. Acoustic and Imagery Semantic Labeling and Fusion of Human-Vehicle Interactions,
SPIE Defense and Security Conference. , . : ,

10/11/2011 11.00 Shirkhodaie, A., Elangovan V.. Context-Based Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,
SPIE Defense and Security Conference. , . : ,

10/11/2011 12.00 Shirkhodaie, A., . Perceptual Semantic Labeling of Human-Vehicle Interactions (HVI),
Second Annual Human and Light Vehicle Detection Workshop. , . : ,

10/11/2011 13.00 Shirkhodaie, A. Semantic Labeling of Human-Vehicle Interactions Via Acoustic Events Characterization and Inference,
Second Annual Human and Light Vehicle Detection Workshop. , . : ,

10/12/2011 19.00 David hall. Challenges in hard and soft fusion: Worth the effort?,
Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011. , . : ,

TOTAL: 77

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

Received		Paper
02/02/2016	43.00	Alexander Nikolaev, Rakesh Nagi, Mohammadreza Samadi. A Subjective Evidence Model for Influence Maximization in Social Networks, OMEGA (03 2016)
08/28/2012	49.00	Gregory Tauer, Rakesh Nagi, Moises Sudit. The Graph Association Problem:Mathematical Models and a Lagrangian Heuristic, Navel Research Logistics (11 2011)
08/29/2012	62.00	Ronald R. Yager. Conditional Approach toPossibility-Probability Fusion, IEEE Transactions on Fuzzy Systems (02 2012)
08/29/2012	68.00	Ronald R. Yager. Participatory Learning of Propositional Knowledge, IEEE Transactions on Fuzzy Systems (08 2012)
08/29/2012	67.00	Ronald R. Yager. Entailment Principle for Measure-Based Uncertainty, IEEE Transactions on Fuzzy Systems (06 2012)
08/29/2012	66.00	Ronald R. Yager. Set Measure Directed Multi-SourceInformation Fusion, IEEE Transactions on Fuzzy Systems (12 2011)
08/29/2012	65.00	Ronald R. Yager. Dempster–Shafer structures with general measures, International Journal of General Systems (05 2012)
08/29/2012	64.00	Ronald R. Yager. Expansible measures of specificity, International Journal of General Systems (04 2012)
08/29/2012	63.00	Ronald R. Yager. On Z-Valuations Using Zadeh's Z-Numbers, International Journal of Intelligent Systems (07 2012)
08/29/2013	75.00	Geoff Gross, Rakesh Nagi, Kedar Sambhoos. A fuzzy graph matching approach in intelligence analysisand maintenance of continuous situational awareness, Information Fusion (10 2011)
08/31/2011	1.00	Gross, G., Jenkins, M., Bisantz, A. M., Nagi, R. Towards Context Aware Data Fusion: Modeling and Integration of Situationally Qualified Human Observations into a Fusion Process for Intelligence Analysis, IEEE Transaction on Systems, Man and Cybernetics: Systems and Humans (05 2011)

TOTAL: **11**

Number of Manuscripts:

Books

Received	Book
08/29/2012 55.00	David L. Hall. Distributed Data Fusion for Network-Centric Operations - Perspectives onDistributed Data Fusion, Boca Raton, FL : CRC Press, (11 2012)
08/29/2012 56.00	Jeff Rimland. Distributed Data Fusion for Network-Centric Operations - Service-Oriented Architecture for Human-Centric Information Fusion, Boca Raton, FL : CRC Press, (11 2012)
08/29/2012 59.00	David L. Hall. Distributed Sensor Networks - The Emergence of Human-Centric Information Fusion, Boca Raton, FL: CRC Press, (09 2012)
TOTAL:	**3**

Received	Book Chapter
08/29/2014 11.00	Stuart C. Shapiro, Daniel R. Schlegel. Concurrent Reasoning with Inference Graphs, Switzerland: Springer International Publishing, (08 2013)
08/30/2013 97.00	David Hall. The Emergence of Human-Centric Information Fusion, Pennsylvania: Chapman and Hall/CRC, (09 2012)
08/30/2013 99.00	Jeffrey Rimland. Service Oriented Architecture for Human Centric Information Fusion, Pennsylvania: CRC Press, (08 2012)
08/30/2014 28.00	Ronald R. Yager. Measure Inputs to Fuzzy Rules, Switzerland: Springer International Publishing, (07 2014)
TOTAL:	**4**

Patents Submitted

Patents Awarded

Awards

1. Jenkins, M., Bisantz, A., Llinas, J., and Nagi, R. (2013). Investigating and improving network visualizations' effectiveness at supporting
human sensemaking tasks. In Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, San Diego, CA, October
2013.
Winner, Human Factors and Ergonomics Society Cognitive Engineering and Decision-making Technical Group Best Student paper, 2013.
2. Gross, G., M. P. Jenkins, S. Lacey, A. M. Bisantz, & R. Nagi, Towards context-aware data fusion: Evaluating the benefits of integrating
situationally qualified human observations into fusion processes.
Winner of the 2012 University at Buffalo ISE Graduate Research Competition, March, 2012;
2nd place at the 2012 University at Buffalo School of Engineering and Applied Sciences Research Competition, April 26, 2012

Graduate Students

NAME	PERCENT_SUPPORTED	Discipline
Emily Catherman	0.15	
Brandon Journey	0.25	
Shad Stud	0.25	
Jamal Hasan	0.33	
Anthony Baker	0.33	
Adriann N. Wilson	0.33	
Daniel Scobey	0.33	
Ramon Gonzalez	0.50	
Mark Thelen	0.25	
Diarra Fall	0.25	
Ayeke Tegegne	0.25	
Brent Warner	0.25	
David Potter	0.25	
Pedro Tavares	0.50	
Dong Chen	0.08	
Rob Grace	0.50	
Anushra Godbole	0.25	
Matt Lesniewski	0.50	
Aditya Sridara	0.25	
Jeff Rimland	0.50	
Geoff Gross	0.50	
Ketan Date	0.50	
Dan Schegel	0.50	
Michael Jenkins	0.50	
Sushant Khopkar	0.00	
Hossein Matin	0.00	
Michael Stearns	0.10	
Yao Li	0.50	
Vinayak Elangovan	0.50	
Amjad Alkilani	0.50	
Mohammad Habibi	0.00	
Jerry Sweafford	0.25	
Bashir Alsaidi	0.25	
Vinod Bandaru	0.25	
Michael Kandefer	0.50	
Michael Prentice	0.50	
Paul Bunter	0.00	
Judith Tiferes-Wang	0.25	
David Lavergne	0.50	
Dana Kerker	0.25	
Hiroto Kaku	0.25	
Alireza Farasat	0.25	
Megan Hannigan	0.50	
Deven McMaster	0.50	
FTE Equivalent:	**14.15**	
Total Number:	**44**	

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Kedar Sambhoos	0.40
Geoff Gross	0.40
FTE Equivalent:	**0.80**
Total Number:	**2**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Rakesh Nagi	0.14	
Moises Sudit	0.14	
James Llinas	0.14	
Ann Bisantz	0.14	
Stuart Shapiro	0.14	
Alexander Nikolaev	0.05	
Ronald Yager	0.20	
Amir Shirkhodaie	0.10	
David Hall	0.00	
Michael McNeese	0.03	
Jake Graham	0.10	
Richard Tutwiler	0.10	
Jason Corso	0.05	
G. Cai	0.05	
S. Shafer	0.10	
FTE Equivalent:	**1.48**	
Total Number:	**15**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	Discipline
Jon Mclellen	0.25	
Jessica Eisenhauer	0.25	
James Dobler	0.25	
Jennifer Kearns	0.25	
Alyssa McClure	0.25	
Frank Mollica	0.25	
Throsby Wells	0.25	
Shanney Lacey	0.25	
Georgia Cruz	0.25	
FTE Equivalent:	**2.25**	
Total Number:	**9**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

Names of Personnel receiving masters degrees

NAME
Jerry Sweafford
Bashir Alsaidi
Biniyam Chaka
Fatemeh Vaziriborog
Vinod Bandaru
Michael Kandefer
Michael Prentice
Megan Hannigan
Deven McMaster
Hiroto Kaku
Spoorthi Rao Nimmala

Total Number: **11**

Names of personnel receiving PHDs

NAME
Greg Tauer
Geoff Gross
Haroun Rababaah
Amjad Alkilani
Mohammad Habibi
Vinayak Elangovan
Jeff Rimland
Michael Jenkins
Daniel R. Schlegel

Total Number: **9**

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress

Technology Transfer

Entire Soft Fusion Processing Pipeline (software) that embodies the algorithmic and basic research advances was transitioned to the Army Research Labs in Aberdeen Proving Grounds, MD.

A large number of transitions (see report) for the SYNCOIN Dataset were made (practically al forces, universities and for-profit enterprises).

Tractor, the Natural Language Processing System was also transitioned to
• Jim Hendler (for Army Network Science CTA project)
• Mark A. Thomas, USA CIV (US)
• Gabor (Gabe) Schmera, SPAWAR
• Conversations with John Kelly, Model Software Corp.
• A Clojure library for parsing XML files created by GATE available at:
https://github.com/digitalneoplasm/gate.data.xml
• CSNePS available at: https://github.com/SNePS/CSNePS
• I2WD
• ARL APG

A number of components of the soft fusion pipeline made it to an IARPA project through CUBRC.

FINAL REPORT

SCIENTIFIC PROGRESS SECTION

Army Research Office Multidisciplinary University Research Initiative (MURI) Grant on

Unified Research on Network-based Hard/Soft Information Fusion

Period June, 2009 through July 31, 2014

Drs. Rakesh Nagi* (PI) and Moises Sudit (Co-PI)

University at Buffalo

* Also at University at Illinois at Urbana Champaign

nagi@illinois.edu, sudit@buffalo.edu

1. Abstract and Section Organization

1.1 Abstract

The University at Buffalo (UB) Center for Multisource Information Fusion (CMIF) and its partners including the Pennsylvania State University (PSU), Iona College (Iona), Tennessee State University (TSU) and University of Illinois at Urbana-Champaign (UIUC) have conducted research to develop a generalized framework, mathematical techniques, and test and evaluation methods to address the ingestion and harmonized fusion of Hard and Soft information in a distributed (networked) Level 1 and Level 2 data fusion environment. This research activity is supported by a Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) for "Unified Research on Network-based Hard/Soft Information Fusion," issued by the US Army Research Office (ARO) under the program management of Dr. John Lavery and recently Dr. Joe Myers. This report provides a summary of the five years of progress. The primary Research Thrusts addressed are framed around the major functional components of the JDL Fusion Process; these include:

1. Source Characterization of Soft Data input streams including; human observation.—direct, indirect, open source inputs, linguistic framing, and text processing
2. Common Referencing and Alignment of Hard and Soft Data, especially strategies and methods for meta-data generation for Hard-Soft data normalization
3. Generalized Data Association Strategies and Algorithms for Hard and Soft Data
4. Robust Estimation Methods that exploit associated Hard and Soft Data
5. Dynamic Network-based Effects on Hard-Soft Data Fusion Architectures and Methods
6. Test and Evaluation Methodology Development to include Human-in-the-Loop
7. Extensibility, Adaptability, and Robustness Assessment
8. Fusion Process Framework
9. Technology Concept of Employment

This program is a five-year effort and considered distinctive in being a major academic thrust into the complexities of the hard and soft fusion problem. During the five years progress has been made in the following areas:

- Development and refinement of overall system concept for human-centered information fusion and information processing architecture.
- Development of a test and evaluation approach involving an evolutionary approach that proceeds from "truthed" synthetic hard and soft data to human in the loop campus based experiments.
- Creation, refinement and analysis of a COIN inspired synthetic data set involving both hard and soft data
- Hard sensor data fusion (including continued development of algorithms for fusion of hard sensor data including; 2-D/3-D video data and 3-D Flash LIDAR and selected collection of augmented data to demonstrate object classification)
- Human computer interaction for improved sense-making including; scenario development, meta-data generation and refinement of SYNCOIN data, study and analysis of requisite cognitive tasks and associated workload for sense-making, visualization for distributed sense-making study, and prototype implementation of human-computer interaction to support situation analysis of hard/soft data
- Development of a supporting infrastructure for integration of emerging algorithms
- Development of a taxonomy for characterizing the human as observer (source characterization), and uncertainty characterization under environmental and observer characteristics
- Development of Tractor for processing text messages in multiple stages and common referencing; evaluated syntactic and semantic processing techniques and selected GATE (General Architecture for Text Engineering) for syntactic processing and FrameNet for a semantic processing database
- Refinement of soft data association prototype which extends the traditional hypothesis generation-hypothesis evaluation-hypothesis selection paradigm for fusion of soft data and utilizes a data graph association process
- Development of parallel data association algorithms (Hadoop/HBase/map reduce) for handling large scale data.
- Implementation of state estimation algorithm using "dirty" (stochastic) graph matching techniques
- Link Analysis using Hadoop/map-reduce
- Development of robust hard sensor fusion techniques for characterization and semantic annotation of social network activities
- Conduction of calibrated experiments for testing and evaluation of new fusion techniques
- Transition of soft processing stream and hard sensor processing techniques to ARL.

- Developed new methods for representing uncertainty in soft data to support common referencing; explored conditional approach to possibility-probability fusion, imprecise uncertainty measure using belief structures, and aggregation operators to link types of monotonic set measures for uncertainty.
- Developed accurate and efficient fusion algorithms for a set of random attributed graphs
- Designed and implemented an infrastructure to support distributed information fusion using communication methods and protocols, extensions of Service Oriented Architecture (SOA) and Message Oriented Middleware (MOM) paradigms, optimized information flow and tasking, complex event processing, and utilization of community standard data representations.
- Knowledge Discovery in Group Activities through Sequential Observation Analysis.
- Basic research in uncertainty representation.
- Visual analytics and cognitive assessment.
- Social Network Analysis for High Value Individual Identification.
- Comprehensive test and evaluation methodology for the Hard-Soft Fusion Architecture.

The project team has been very active in connecting with U. S. Army and Department of Defense end users to assist in understanding the overall problem and guiding development of test and evaluation and operational concepts. In addition, the team has worked with key industrial partners to obtain information that relates to current practices and related programs. The research group has also pursued transition opportunities with other government organizations for some of the mature components.

The remainder of this document provides a perspective on the problem and solution space and information from each team member on their accomplishments, project statistics, and publications.

1.2 Section Organization

This narrative is the Scientific Progress portion of our Final Report to ARO. Our MURI Team comprises the University at Buffalo as the lead university, and University of Illinois at Urbana-Champaign, Penn State University, Tennessee State University, and IONA College as our Team-mates. This major section provides our high-level perspectives of the challenges of this program and our Scientific Progress as described for each Team member. While the order does not reflect any ordering criteria, the following report ordering for this report Scientific Progress section is the same as for the rest of the Continuation Sheet, with the writings ordered in the following way:

A. University at Buffalo
B. University of Illinois
C. Penn State University
D. Tennessee State University
E. IONA College

For the other portions of the continuation sheets of SF298, since there are many sections and subsections involved, we have major sectional categories to allow the reader ease of navigation.

Table of Contents

2. Program Overview Narrative

2.1 Background and Program Perspective

This ARO research grant is addressing some of the most modern and challenging problems in information processing that face the US Army in its current worldwide operations. It is addressing the overarching issue of automatically or semi-automatically forming the best possible estimates of situational state estimates by Information Fusion operations on a plethora of disparate and uncertain observational and contextual data and information sources streaming in from a dynamically changing operational environment. The complexity of forming such estimates is compounded by the combination of data that is uncertain, ambiguous, and of mixed reliability coupled with the operational problem environment that involves insurgency within a foreign population.

Insurgencies and the methods of Counter-Insurgency ("COIN") operations are extraordinarily complex environments to deal with and even to define. From Army FM3-24 on Counterinsurgency, we have a definition of insurgency as: "Joint doctrine defines an insurgency as an organized movement aimed at the overthrow of a constituted government through the use of subversion and armed conflict." From the same source, we have the definition "Counterinsurgency is military, paramilitary, political, economic, psychological, and civic actions taken by a government to defeat insurgency." Thus, these conflicts do not involve known, uniformed adversaries, and have very high collateral damage considerations since the conflicts occur within neutral populations. Course of action choices are both highly varied, involving all the factors just mentioned but at the same time is highly constrained. The MURI Problem Domain is considered to be the problem of Small-scale COIN insurgency. In Small-scale insurgencies, belligerent groups have established some size, are developing tactics techniques and procedures, and are causing hostile and possibly lethal events. These groups however are still quite covert and operate very carefully; their leadership and organizational structures and their insurgency-related goals and objectives are still not well understood.

Considering the Small-scale COIN problem, the requirements for Information Fusion (IF) are to estimate the "essential elements of information (EEIs)" for this sub-problem space of COIN, in support of corresponding military or other possible courses of action. The framework for research planning for the MURI has thus been developed around a "requirements relevant" but not "requirements-driven" approach to the prototyping of a Hard-Soft IF process; that is, this research program has no operationally-specific Army requirements specification or specific application paradigm. The positive side of this is that the research will not yield a "point design." For some specific operational application, it should yield an architecture that is flexible to new data sources. However, there is in fact some risk of non-applicability. To deal with this in part the program includes a task to examine scalability and robustness of developed solution strategies. It is intended that these planning aspects be worked in conjunction with the ARO, and other Army organizations as we have already begun in the base funding period.

Another critical research strategy choice is that, based on extensive analysis, we should preferably have chosen an inductive, learning/discovery-based approach regarding the development of insight for a dynamic COIN problem. Modern literature shows that the ability to effectively model human group dynamics and relationships remains a very challenging problem

and that only very limited capability exists. We have largely focused on deductive or model based discovery in the project because it forms the basis of inductive and abductive approaches.

In particular for the Soft Data Fusion problem, the UB team has chosen graph-based methods as an inferencing framework, wherein the soft data are associated and batched into an evolving, accumulating “Data Graph” representing cumulating situational evidence, and using this Data Graph and analyst formed queries that can also be represented as graphs (“Target Graphs”), state of the art methods developed at CMIF are used for Graph Matching to yield inferred assertions, supporting an adaptive analyst learning process. The operational focus here is on human social network type inquiries. On the Hard Data Fusion side, PSU and TSU are combining to use multispectral Hard Data of various modes (e.g. imagery, acoustic, video) to focus on the issue of human-vehicle behaviors and relationships, since vehicles and their use in various ways have proven critical in COIN type operational problems.

Using the Uncertainty knowledge base developed by UB in conjunction with IONA college provides a framework for uncertainty alignment in the graph based soft fusion process. This uncertainty knowledge is also part of the hard-soft fusion framework developed by UB. In hard-soft fusion the kinematic tracks with from hard data (PSU) along with acoustic signatures (TSU) are merged with location information for various entities (persons, locations etc.).

3. Accomplishments and Narratives of Research Efforts

3.1 University at Buffalo

List of papers submitted or published:

- Papers Published in Peer-reviewed Journals

1. G. Tauer, R. Nagi, and M. Sudit, “The graph association problem: mathematical models and a Lagrangian heuristic,” Naval Research Logistics, vol. 60, pp. 251-268, April 2013
2. G. Tauer, and R. Nagi, “A Map-Reduce Lagrangian heuristic for multidimensional assignment problems with decomposable costs,” Parallel Computing, 39(11), pp. 653-658, November 2013.
3. Jenkins, M., Gross, G., Bisantz, A. and Nagi, R. "Towards Context Aware Data Fusion: Modeling and Integration of Situationally Qualified Human Observations into a Fusion Process for Intelligence Analysis," *Information Fusion,* January 2015, Vol. 21, pp. 130-144.
4. Gross, G.A., Nagi, R. and Sambhoos, K. "A Fuzzy Graph Matching Approach in Intelligence Analysis and Maintenance of Continuous Situational Awareness," *Information Fusion*, July 2014, Vol. 18, pp. 43-61.
5. Date, K. and Nagi, R. "A GPU Accelerated Hungarian Algorithm for the Linear Assignment Problem," submitted to *Parallel Computing*, June 2014.
6. Gross, G.A. and Nagi, R. "Precedence Tree Guided Search for the Efficient Identification of Multiple Situations of Interest – AND/OR Graph Matching," submitted to *Information Fusion*, June 2014.
7. Michael Stearns, & Alexander Nikolaev, "Modeling and Recognition of Complex Temporal Events in SmartHome Environment", submitted to *IEEE Pervasive Computing*, in the second round of review.

8. Michael Stearns, & Alexander Nikolaev, "A Random Graph Entropy-Based Approach for Complex Activity Recognition", submitted to *IIE Transactions*, in the first round of review.

- Papers published in peer-reviewed conference proceedings

1. K. Date, G. A. Gross, S. Khopkar, R. Nagi, K. Sambhoos. 2013. "Data association and graph analytical processing of hard and soft intelligence data," Proceedings of the 16th International Conference on Information Fusion (Fusion 2013), Istanbul, Turkey, 09-12 July 2013.
2. Llinas, J., Reexamining Information Fusion--Decision Making Inter-dependencies, Presented at IEEE CogSIMA (Cognitive Situation Management) Conference, San Antonio, TX, Mar 2014
3. Llinas, J., A survey of automated methods for sensemaking support, Presented at the SPIE Next-Generation Analytics Conference (part of SPIE Defense and Security Conf), Baltimore, MD, April 2014.
4. Llinas, J., Challenges in Information Fusion Technology Capabilities for Modern Intelligence and Security Problems, presented at the European Intelligence and Security Informatics Conference (EISIC) 2013, August 12-14, 2013 Uppsala, Sweden.
5. G. Cai, G. Gross, J. Llinas and D. Hall, *A Visual Analytic Framework for Data Fusion in Investigative Intelligence,* 2014 SPIE DSS – Next Generation Analyst 2, Baltimore, MD, May 2014.
6. K. Date, G. A. Gross, R. Nagi. 2014. Test and Evaluation of Data Association Algorithms in Hard+Soft Data Fusion. Proceedings of the 17th International Conference on Information Fusion (Fusion 2014), Salamanca, Spain.
7. G. A. Gross, K. Date, D. R. Schlegel, J. Corso, J. Llinas, R. Nagi, S. Shapiro. 2014. Systemic Test and Evaluation of a Hard+Soft Information Fusion Framework. Proceedings of the 17th International Conference on Information Fusion (Fusion 2014), Salamanca, Spain.
8. Kerker, D., Jenkins, M.P, Gross, G.A., Bisantz, A. and Nagi, R. "Visual Estimation of Human Attributes: An empirical study of context-dependent human observation capabilities," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, San Antonio, TX, 3-6 March 2014.
9. Gross, G., Nagi, R., Sambhoos, K., Schlegel, D., Shapiro, S. and Tauer, G. "Towards Hard+Soft Data Fusion: Processing Architecture and Implementation for the Joint Fusion and Analysis of Hard and Soft Intelligence Data," *15th International Conference on Information Fusion*, Singapore, 9-12 July 2012.
10. Blasch, E., Costa, P.C.G., Laskey, K.B., Stampouli, D., Ng, G.W., Schubert, J., Nagi, R., and Valin, P. "Issues of Uncertainty Analysis in High-Level Information Fusion," *15th International Conference on Information Fusion*, Singapore, 9-12 July 2012.
11. McConky, K., Nagi, R., Sudit, M. and Hughes, W. "Improving Event Co-reference By Context Extraction and Dynamic Feature Weighting," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, New Orleans, LA, 6-8 March 2012.

12. Gross, G., Nagi, R. and Sambhoos, K. "Continuous Preservation of Situational Awareness through Incremental/Stochastic Graphical Methods," *14th International Conference on Information Fusion*, Chicago, IL, 26-29 July 2011.
13. McMaster, D., Nagi, R. and Sambhoos, K. "Temporal Alignment in Soft Information Processing," *14th International Conference on Information Fusion*, Chicago, IL, 26-29 July 2011.
14. Jenkins, M.P., Gross, G., Bisantz, A. and Nagi, R. "Towards context-aware hard/soft information fusion: Incorporating situationally qualified human observations into a fusion process for intelligence analysis," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, Miami Beach, FL, 22-24 February, 2011.
15. Gross, G., Nagi, R. and Sambhoos, K. "Soft Information, Dirty Graphs and Uncertainty Representation/Processing for Situation Understanding," *13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July 2010.
16. Llinas, J., Nagi, R., Hall, D. and Lavery, J. "A Multi-disciplinary University Research Initiative in Hard and Soft Information Fusion: Overview, Research Strategies and Initial Results," *13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July 2010.
17. Gross, G., Nagi, R. and Sambhoos, K. "Situation Assessment: Uncertainty Representation in Inexact Graph Matching," *16th Industrial Engineering Research Conference*, Cancun MX, June 2010.
18. Llinas, J., Nagi, R., Duff, D., Patel, M. and Walsh, D. "Framing and Defining New Fusion Strategies and Advanced Analytics for Relation-driven Problem Environments," *2012 Military Sensing Symposia, National Symposium on Sensor and Data Fusion (NSSDF),* Washington, DC, October 2012.
19. Juan Gómez-Romero, Jesús Garcia, Michael Kandefer, James Llinas, Jose Manuel Molina, Miguel Angel Patricio, Michael Prentice, and Stuart C. Shapiro, "Strategies and Techniques for Use and Exploitation of Contextual Information in High-Level Fusion Architectures", *Proceedings of the 13th International Conference on Information Fusion (Fusion2010)*, Edinburgh, UK. July 2010., TH1.7.3, 8 pages, unpaginated.
20. Michael Prentice, Michael Kandefer, and Stuart C. Shapiro, "Tractor: A Framework for Soft Information Fusion", *Proceedings of the 13th International Conference on Information Fusion (Fusion2010)*, Edinburgh, UK. July 2010., Th3.2.2, 8 pages, unpaginated.
21. Stuart C. Shapiro, The Jobs Puzzle: A Challenge for Logical Expressibility and Automated Reasoning. In E. Davis, P. Doherty, and E. Erdem, Eds., *Logical Formalizations of Commonsense Reasoning: Papers from the AAAI Spring Symposium*, Technical Report SS-11-06, AAAI Press, Menlo Park, CA, 2011, 96-102.
22. Michael Kandefer and Stuart C. Shapiro, Evaluating Spreading Activation for Soft Information Fusion, *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, 2011.
23. Michael Prentice and Stuart C. Shapiro, Using Propositional Graphs for Soft Information Fusion, *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, 2011.

24. Daniel R. Schlegel and Stuart C. Shapiro, Visually Interacting with a Knowledge Base Using Frames, Logic, and Propositional Graphs, *Second IJCAI International Workshop on Graph Structures for Knowledge Representation and Reasoning*, 2011.
25. Stuart C. Shapiro, and Daniel R. Schlegel, Natural Language Understanding for Soft Information Fusion, *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013),* IFIP, July, 2013, unpaginated, 9 pages.
26. Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs (student abstract), *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, AAAI Press/The MIT Press, Menlo Park, CA, 2013, 1637-1638.
27. Daniel R. Schlegel, Concurrent Inference Graphs (Doctoral Consortium abstract), *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, AAAI Press/The MIT Press, Menlo Park, CA, 2013, 1680-1681.
28. Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs. In *Working Notes of the 3rd International IJCAI Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR@IJCAI 2013)*, 2013, unpaginated, 9 pages.
29. Daniel R. Schlegel and Stuart C. Shapiro, Inference Graphs: A Roadmap. In Matthew Klenk and John Laird, Eds. *Proceedings of the Second Annual Conference on Advances in Cognitive Systems*, 2013 Poster Collection, December, 2013, 217-234.
30. Daniel R. Schlegel and Stuart C. Shapiro, Inference Graphs: A New Kind of Hybrid Reasoning System (student abstract), *Proceedings of the Twenty-Eighth Conference on Artificial Intelligence (AAAI-14),* AAAI Press/The MIT Press, Menlo Park, CA, 2014, 3134-3135.
31. Schlegel, D.R. and Shapiro, S.C., Inference Graphs: A New Kind of Hybrid Reasoning System, *Proceedings of the Cognitive Computing for Augmented Human Intelligence Workshop at AAAI-14 (CGAHI@AAAI-14),* 2014, 4 pages, 38-41.
32. Daniel R. Schlegel and Stuart C. Shapiro, The `Ah Ha!' Moment : When Possible, Answering the Currently Unanswerable using Focused Reasoning. In P. Bello, M. Guarini, M. McShane, & B. Scassellati, Eds., *Proceedings of the 36th Annual Conference of the Cognitive Science Society (COGSCI 2014)*, Cognitive Science Society, Austin, TX, 2014, 1371-1376.
33. Michael Stearns, Alexander Nikolaev, Sue Kase, & Kirk Ogaard, , 1371-1376.e Science Society (COGSCI 2014)o, M. Guarini, M. McShane, & B. Scassell*IIE ISERC 2013 Conference*, San Juan, Puerto Rico, May 2013.
34. Hannigan, M., McMaster, D., and Llinas, J., Data Association and Soft Data Streams, Proc of the Intl Conf on Information Fusion, July 2010, Edinburgh, UK

- Papers presented at peer-reviewed conferences

1. K. Date, G. A. Gross, S. Khopkar, R. Nagi, K. Sambhoos. 2013. “Data association and graph analytical processing of hard and soft intelligence data,” 16th International Conference on Information Fusion (Fusion 2013), Istanbul, Turkey.
2. K. Date, G. A. Gross, R. Nagi. 2014. “Test and Evaluation of Data Association Algorithms in Hard+Soft Data Fusion,” 17th International Conference on Information Fusion (Fusion 2014), Salamanca, Spain.

3. G. A. Gross, K. Date, D. R. Schlegel, J. Corso, J. Llinas, R. Nagi, S. Shapiro. 2014. “Systemic Test and Evaluation of a Hard+Soft Information Fusion Framework,” 17th International Conference on Information Fusion (Fusion 2014), Salamanca, Spain.
4. Kerker, D., Jenkins, M.P, Gross, G.A., Bisantz, A. and Nagi, R. "Visual Estimation of Human Attributes: An empirical study of context-dependent human observation capabilities," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, San Antonio, TX, 3-6 March 2014.
5. Gross, G., Nagi, R., Sambhoos, K., Schlegel, D., Shapiro, S. and Tauer, G. "Towards Hard+Soft Data Fusion: Processing Architecture and Implementation for the Joint Fusion and Analysis of Hard and Soft Intelligence Data," *15th International Conference on Information Fusion*, Singapore, 9-12 July 2012.
6. Blasch, E., Costa, P.C.G., Laskey, K.B., Stampouli, D., Ng, G.W., Schubert, J., Nagi, R., and Valin, P. "Issues of Uncertainty Analysis in High-Level Information Fusion," *15th International Conference on Information Fusion*, Singapore, 9-12 July 2012.
7. McConky, K., Nagi, R., Sudit, M. and Hughes, W. "Improving Event Co-reference By Context Extraction and Dynamic Feature Weighting," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, New Orleans, LA, 6-8 March 2012.
8. Jenkins, M.P.; Bisantz, A.M., “Identification of Human-Interaction Touch Points for Intelligence Analysis Information Fusion Systems,” FUSION 2011, 14th International Conference on Information Fusion, Special Session on Hard/Soft Information Fusion: New Data Sets and Innovative Architectures. 8 pp., 5-8 July 2011.
9. Gross, G., Nagi, R. and Sambhoos, K. "Continuous Preservation of Situational Awareness through Incremental/Stochastic Graphical Methods," *14th International Conference on Information Fusion*, Chicago, IL, 26-29 July 2011.
10. McMaster, D., Nagi, R. and Sambhoos, K. "Temporal Alignment in Soft Information Processing," *14th International Conference on Information Fusion*, Chicago, IL, 26-29 July 2011.
11. Jenkins, M.P., Gross, G., Bisantz, A. and Nagi, R. "Towards context-aware hard/soft information fusion: Incorporating situationally qualified human observations into a fusion process for intelligence analysis," *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, Miami Beach, FL, 22-24 February, 2011.
12. Gross, G., Nagi, R. and Sambhoos, K. "Soft Information, Dirty Graphs and Uncertainty Representation/Processing for Situation Understanding," *13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July 2010.
13. Llinas, J., Nagi, R., Hall, D. and Lavery, J. "A Multi-disciplinary University Research Initiative in Hard and Soft Information Fusion: Overview, Research Strategies and Initial Results," *13th International Conference on Information Fusion*, Edinburgh, Scotland, 26-29 July 2010.
14. Gross, G., Nagi, R. and Sambhoos, K. "Situation Assessment: Uncertainty Representation in Inexact Graph Matching," *16th Industrial Engineering Research Conference*, Cancun MX, June 2010.
15. Llinas, J., Nagi, R., Duff, D., Patel, M. and Walsh, D. "Framing and Defining New Fusion Strategies and Advanced Analytics for Relation-driven Problem Environments,"

2012 Military Sensing Symposia, National Symposium on Sensor and Data Fusion (NSSDF), Washington, DC, October 2012.

- Other presentations
  1. Llinas, J., Gross, G. and Nagi, R., "Challenges of and Approaches to Hard-Soft Information Fusion," Tutorial, Cognitive Situation Management (CogSIMA 2014), March 2014, San Antonio, TX.
  2. Ogaard, K., Roy, H., Kase, S., Nagi, R., Sambhoos, K. and Sudit, M. "Discovering Patterns in Social Networks with Graph Matching Algorithms," *2013 International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction (SBP 2013),* Washington, DC, April 2013.
  3. Llinas, J. and Nagi, R. "Tutorial: Challenges and Approaches to Hard-Soft Information Fusion," *2012 Military Sensing Symposia, National Symposium on Sensor and Data Fusion (NSSDF),* Washington, DC, October 2012.
  4. Nagi, R. "Information Fusion and Intelligence Analysis with Hard and Soft Data: Industrial and Systems Engineering Challenges and Opportunities," Department of Integrated Systems Engineering, Ohio State University, August 2014.
  5. Nagi, R. "[Big Data] Motivation and Big Graph Challenges," Big Data Research Workshop by Computational Science and Engineering, University of Illinois Urbana-Champaign, May 2014.
  6. Nagi, R. "Fusion of Hard and Soft Information and The Graph Association Problem," Computational Science and Engineering, University of Illinois Urbana-Champaign, February 2014.
  7. Stuart C. Shapiro, Tractor: Toward Deep Understanding of Short Intelligence Messages, presented to the UB Center for Cognitive Science, March 28, 2012.
  8. Stuart C. Shapiro, Tractor: Toward Deep Understanding of Short Intelligence Messages, presented to Milan Patel, 2WD/A2SF , June 19, 2012.

b) Manuscripts

1. G. Tauer, K. Date, R. Nagi, and M. Sudit, "An incremental graph-partitioning algorithm for entity resolution," Under Revision. To be submitted to Transactions on Knowledge Discovery from Data.
2. LaVergne, D., Tiferes, J., Jenkins, M., Gross, G., and Bisantz, A. M. Linguistic Descriptors of Human Attributes. Submitted to IEEE Transactions on Human-machine systems, August, 2014.
3. Stuart C. Shapiro, Michael Prentice, and Daniel R. Schlegel, Tractor Manual, University at Buffalo, July 18, 2012.
4. Stuart C. Shapiro, A Grading Rubric for Soft Information Understanding, Department of Computer Science and Engineering, University at Buffalo, December 13, 2013.

c) Books and Book Chapters

1. Jenkins, M., Bisantz, A. M., and Pfautz, J. (2012) Human Engineering Factors in Distributed and Net Centric Fusion Systems. In D. Hall, C. Chong, J. Llinas, and M.

Liggins II, Eds., Distributed Data Fusion for Network-centric Operations. Taylor and Francis, pp 409 – 434.

2. Daniel R. Schlegel and Stuart C. Shapiro, Visually Interacting with a Knowledge Base Using Frames, Logic, and Propositional Graphs. In Madalina Croitoru, Sebastian Rudolph, Nic Wilson, John Howse and Olivier Corby, Eds., *Graph Structures for Knowledge Representation and Reasoning, Lecture Notes in Artificial Intelligence 7205*, Springer-Verlag, Berlin, 2012, 188-207.
3. Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs. In Madalina Croitoru, Sebastian Rudolph, Stefan Woltran, and Christophe Gonzales, Eds., *Graph Structures for Knowledge Representation and Reasoning, Lecture Notes in Artificial Intelligence 8323*, Springer International Publishing, Switzerland, 2014, 138-164.
   DOI: 10.1007/978-3-319-04534-4_10
4. Stuart C. Shapiro and Daniel R. Schlegel, Natural Language Understanding for Information Fusion. In Galina Rogova and Peter Scott, Eds., *Fusion Methodologies in Crisis Management Higher Level Fusion and Decision Making for Crisis Management*, Springer S. A., in press.
5. Michael Kandefer and Stuart C. Shapiro, Context Relevance for Text Analysis and Enhancement --- Soft Information Fusion. In Lauro Snidaro, Jesus Garcia, Erik Blasch and James Llinas, Eds., *Boosting Real World Performance with Domain Knowledge*, Springer S. A., in preparation.

d) Theses and dissertations

1. G. Tauer (2012). “Data Association on Large Quantities of Complex Data.” PhD dissertation in Department of Industrial and Systems Engineering, The State University of New York at Buffalo, August 2012.
2. Katie McConky; graduated 12/12. Dissertation title: "Applications of Location Similarity Measures and Conceptual Spaces to Event Coreference and Classification." (co-advised by Rakesh Nagi and Moises Sudit.) Currently: CUBRC, NY.
3. Geoff Gross; graduated 5/13. Dissertation title: "Graph Analytic Techniques in Uncertain Environments: Graph Matching and Link Analysis." Advised by Rakesh Nagi, Currently: Research Associate, Center for Multisource Information Fusion, University at Buffalo (SUNY).
4. Jenkins, M. Towards a lexicon of visualization design templates: Supporting sensemaking with enhanced network visualizations. University at Buffalo, State University of NY at Buffalo, 2013, 646 pages.
5. Daniel R. Schlegel, Concurrent Inference Graphs, PhD Dissertation, August, 2014.
6. Ketan Date; expected graduation 9/15. Tentative dissertation title: "Assignment and Association Problems." University of Illinois at Urbana-Champaign.

e) Data Sets

1. Jenkins, M. P.; Bisantz, A.; Llinas, J.; Nagi, R. (2014). MAVERICK Synthetic Murder Mystery Dataset (version 1.0) [data files and ground truth]. Retrieved from the University

of Buffalo, State University of New York (SUNY) Institutional Repository: http://hdl.handle.net/10477/24359

Honors and Awards –

Jenkins, M., Bisantz, A., Llinas, J., and Nagi, R. (2013). Investigating and improving network visualizations' effectiveness at supporting human sensemaking tasks. In Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting, San Diego, CA, October 2013.

1. Winner, Human Factors and Ergonomics Society Cognitive Engineering and Decision-making Technical Group Best Student paper, 2013.

Gross, G., M. P. Jenkins, S. Lacey, A. M. Bisantz, & R. Nagi, Towards context-aware data fusion: Evaluating the benefits of integrating situationally qualified human observations into fusion processes

1. Winner of the 2012 University at Buffalo ISE Graduate Research Competition, March, 2012;
2. 2nd place at the 2012 University at Buffalo School of Engineering and Applied Sciences Research Competition, April 26, 2012;

Titles of Patents disclosed during the reporting period –

Patents awarded during the reporting period –

Graduate Students

Name	**Per Cent Supported**
Michael Kandefer	50%
Michael Prentice	50%
Daniel R. Schlegel	50%
Paul Bunter	Volunteer

Michael Jenkins	50%
Judith Tiferes-Wang	50%
David Lavergne	50%
Dana Kerker	50%
Hiroto Kaku	50%
Geoff Gross	50%
Greg Tauer	50%
Ketan Date	50%
Hossein Nick Zinat Matin	50%
Gang Chen	50%
Alireza Farasat	50%
Michael Stearns	50%
Megan Hannigan	50%
Deven McMaster	50%
Total Number:	18

Post Doctorates

Name	**Per Cent Supported**
Geoff Gross	%
Kedar Sambhoos	%
Total Number:	2

Faculty

Name	Percent Supported
Ann Bisantz	~10%
Alexander Nikolaev	~5%
Jason Corso	~5%
Stuart C. Shapiro	~10%
Moises Sudit	~10%
James Llinas	~10%
Rakesh Nagi	~10%
Total Number:	7

Undergraduate Students

Name	Percent Supported
Jon Mclellen	~25%
Jessica Eisenhauer	~25%
James Dobler	~25%
Jennifer Kearns	~25%
Alyssa McClure	~25%
Frank Mollica	~25%
Throsby Wells	~25%
Shanney Lacey	25%
Georgia Cruz	25%
Total Number:	**9**

Student Metrics

The number of post-graduates & PhDs funded during this period	2
The number of under-graduates funded during this period	1
The number of undergraduates funded who graduated during this period	1
The number of undergraduates who graduated during this period with a degree in science, mathematics, engineering, or technology fields	1
The number of undergrads who graduated during this period and will continue to pursue a graduate or PhD degree in science, mathematics, engineering or technology fields	0
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0	N/A
Number of graduating undergrads funded by a DoD funded Center of Excellence grant for Education, Research and Engineering	
The number of undergrads who graduated during this period and intend to work for the Department of Defense	
The number of undergraduates who graduated during this period and will receive scholarships or fellowships to further studies in science, mathematics, engineering or technology fields	

Masters Degrees Awarded (5, 3 non-thesis)

- Michael Prentice, MS in CSE, "Tractor: An Architecture for Natural Language Processing"
- Megan Hannigan, 2011 MS in ISE, "Design challenges in developing a soft data association process."

PhDs Awarded (4)

- G. Tauer (2012). Data Association on Large Quantities of Complex Data. PhD dissertation in Department of Industrial and Systems Engineering, The State University of New York at Buffalo, August 2012.
- Geoff Gross (2013). "Graph Analytic Techniques in Uncertain Environments: Graph Matching and Link Analysis." PhD dissertation in Department of Industrial and Systems Engineering, The State University of New York at Buffalo, May 2013.

- Jenkins, M. (2013) Towards a lexicon of visualization design templates: Supporting sensemaking with enhanced network visualizations. University at Buffalo, State University of NY at Buffalo, 2013, 646 pages.
- Daniel R. Schlegel, PhD in CSE, "Concurrent Inference Graphs"

Other Research Staff –

Technology transfer

- Tractor
    - Jim Hendler (for Army Network Science CTA project)
    - Mark A. Thomas, USA CIV (US)
    - Gabor (Gabe) Schmera, SPAWAR
    - Conversations with John Kelly, Model Software Corp.
    - A Clojure library for parsing XML files created by GATE available at: https://github.com/digitalneoplasm/gate.data.xml
    - CSNePS available at: https://github.com/SNePS/CSNePS
- I2WD
- ARL APG

3.1.1 Human Source Characterization and Network Visualization

3.1.1.1 Human Source Characterization

Characterizing human observations in terms of their accuracy and reliability, under different task and environmental conditions, is a key challenge for hard+soft information fusion systems. Our research began by first developing a taxonomy of likely types of human observations for COIN operations and then defining context-driven error characterization models for respective categories. The taxonomy was derived from a categorical analysis of human observations present in the STEF data set as well as observational categories described in COIN operation and US Army field manuals. Sixty-seven categories were identified, including type of observation, method of observation (e.g. visual or auditory), and type of response (e.g. qualitative or quantitative). For instance, one can observe age by either listening to a voice or seeing a person, and express that age as either a number ("21 years") or by using words ("young man"). There is inherent variability in human perceptual and cognitive systems, both within a person (e.g. over time) and across different people, as well as an essentially infinite combination of environmental and task combinations that can impact observations. Developing a complete set of observational error characteristics empirically is therefore impractical.

We performed a focused search of the psychological, perceptual, and judgment literature in order to identify pre-existing empirical results related to human observational errors associated with the 67 identified categories. Based on the availability of data from the literature, and/or utility for processing messages from the STEF data set, four observational categories were selected for further investigation in Year 1: quantitative egocentric distance (distance from an observer to an object), age (quantitative), numerosity (number of objects in a set or group), and time of past events. A process for mapping empirical results drawn from the literature (including

quantitative error estimates and meta-information–such as observational context) to membership functions was developed to support inclusion of these categories of human observation in the fusion algorithms. The use of fuzzy membership functions for situation assessment within data fusion systems is not a new concept. However, membership functions for soft data sources have typically been artificially generated in the past and lack contextual considerations (that, in a real-world application, can drastically change the error characteristics used to generate the membership function). Plans for targeted data collection to fill specific gaps in the empirical data (e.g. qualitative age estimation and exocentric distance) were also developed.

During Year 2, an additional nine categories were selected for investigation and the process developed during Year 1 was applied in order to include the categories in the fusion algorithms. The additional categories were: visual based facial recognition, visual based object dimension estimation (quantitative), visual based gender classification, visual based large scale (>300 people) crowd size estimation (quantitative), recall based duration estimation of past events (quantitative), auditory based voice recognition, visual based pitch estimation (quantitative), and visual/haptic/recall based traversed distance estimation (quantitative).

In Year 3 we focused on validation of the results from Years 1 & 2 by evaluating the benefits of integrating the human observation error characterization models into the fusion system's data association and situation assessment processes via uncertainty alignment. To perform this evaluation, a *synthetic* dataset of observations (i.e. attribute values based on known human estimation capabilities) and an *observed* dataset of actual human observations (i.e. from human participants making observations of simulated insurgency events) were created. The synthetic dataset was generated by leveraging U.S. Census data to populate of list of potential observers as well as a list of potential targets (i.e. observed people). Observations were generated by randomly selecting an observer from the potential observers list and, given their known attributes and the error characterization model for respective categories of observation, adjusting a target person's attributes for the given bias and variance distribution. The benefits of the uncertainty alignment process leveraging the human error characterization models were then evaluated in two ways.

First, to simulate graph matching, observed people that were processed either with or without uncertainty alignment (leveraging the characterization models) were compared against truthed candidates that were manually input into the fusion data graph. Second, to simulate data association, observed people were again processed either with or without uncertainty alignment, but were instead compared against other observed people processed in the same manner. This procedure was replicated to simulate approximately 40,000 attribute estimations with up to five attributes being observed on a single person. Initial results for true person-uncertainty-aligned-observed persons showed (i) a 6.7% increase in similarity comparisons $\leq$ 1% away from top similarity score after uncertainty alignment process, and (ii) a 5.9% increase in similarity comparisons (which are the top ranked similarity score after uncertainty alignment process).

A potential limitation of the synthetic dataset approach, however, is that the generation of observations uses the same observation error characterization models that are leveraged by the uncertainty alignment process. For this reason, the observed dataset was generated to address this limitation as well as to help validate the error characterization models and generate estimated linguistic-to-numerical fuzzy distributions for select categories of observations.

Year 3 also focused on the collection of the observed dataset, which entailed recruiting human participants to provide observations about characteristics of actors (e.g. their height, weight, and age) shown in a video recording of a simulated insurgency scenario (generated by PSU). Fourteen attributes of twelve individuals in two separate simulated insurgency scenario videos were collected by thirty human participants of varying demographics, resulting in the collection of approximately 10,000 attribute estimations. Estimates for the age, height, and weight of a person were provided in both numeric and linguistic forms to support the generation of linguistic-numeric mappings.

This data was analyzed over two studies in Years 4 & 5, one of which (i) attempted to characterize how observers' own characteristics influence their ability to make judgments on those they are observing, and another which (ii) investigated how people produce linguistic descriptions of human physical attributes in order to support the development of algorithms which combine hard and soft data in fusion processes. The first analysis found that observers' own age and height were not strong biasing factors in their estimations of targets' age and height (in contrast to previous studies), though weight may have been. The second analysis examined the feasibility of creating fuzzy membership functions to facilitate qualitative-quantitative mapping stage in order for fusion systems to be able to process linguistically-reported human observations. We found that while people may use a multitude of terms to describe the same attribute values, only relatively few were used with any consistency. This leads us to hypothesize that a controlled-language lexicon may be useful for qualitative-quantitative mapping, but more research is needed.

3.1.1.2 Fusion System Data Graph Visualization

There has been minimal support for visualization designers in terms of guidance for developing effective visualizations based on an understanding of human perceptual and cognitive capabilities, especially within the realm of network visualizations. This is relevant, as the output of many fusion systems–including the one developed in this research effort–are data graphs presented in the form of a network visualization. While our efforts in Years 1 and 2 focused on the development of a network visualization user-interface (to allow basic functionality from a testing and evaluation perspective), developing a user interface focused on supporting hypothesized end-users is a challenging problem that remained out of the scope of this research effort during that time.

In Year 3, focus was finally given to this challenge by characterizing the types of tasks that intelligence analyst typically encounter, based on existing work analysis literature and leveraging that characterization to identify (i) which common tasks the fusion system data graph has potential to support, and (ii) where design enhancements could potentially increase or contribute additional task specific performance benefits. This allowed for potential design enhancements for the data graph network visualization to be generated with features that are hypothesized to better support analyst performance for the set of identified intelligence analysis tasks. To evaluate the hypothesized performance benefits of the enhanced network visualization, as well as to empirically validate the effectiveness of the network visualization format in general at presenting information to analysts, an empirical study was conducted in Year 4.

The experiment involved human participants carrying out analogous intelligence analysis sensemaking tasks (i.e. information foraging, hypothesis generation, and hypothesis evaluation) to compare the effectiveness of various information displays. As access to intelligence analysts

was not feasible, an analogous domain was adopted (a murder mystery *a la* the board game "Clue") and sample datasets were generated so that participants could be recruited from the Buffalo community with minimal restrictions. Multiple enhanced network visualization information displays were developed to support improved sensemaking task performance. From among these designs, two were selected for empirical evaluation against a non-visualization based display (flat "spreadsheet" data tables), and a traditional network visualization (annotated circles and lines which represent nodes and links respectively). The two selected designs were comprised of unique *node* designs (in order to represent multiple attribute and meta-information variables related to the represented *entity*) and a shared *link* design (to represent multiple attribute and meta-information variables related to the represented *relationship*). Figure 1 and Figure 2 show the various graphic dimensions of the node that were mapped to different task-relevant variables for the two enhanced network visualization designs selected. A total of eleven variables were mapped to different graphic dimensions of the nodes. Figure 3 shows the link design that was investigated, which mapped three distinct variables to different graphical dimensions of the link. Finally, Figure 4 and Figure 5 show sample views of the network visualization displays used in the research study with the different node designs and the shared link design instantiated.

The primary purpose of the research effort was to empirically investigate whether or not network visualizations are an effective and efficient means for supporting human sensemaking tasks compared to non-visualization displays. The second goal was to empirically demonstrate that network visualizations designed to support human behavior at different levels of cognitive control (i.e. the enhanced network visualizations) will better support sensemaking tasks than traditional network visualizations or non-visualization displays. Performance was measured by capturing participants' ability to (i) identify relevant information rapidly, accurately, and robustly, (ii) generate multiple feasible competing hypotheses, and (iii) accurately evaluate and rank the relative accuracy of different hypotheses given limited available information.

Results from the empirical studies showed that basic network visualizations failed to offer any performance benefits over non-visualization displays for both information foraging and sensemaking (i.e., hypothesis generation and evaluation) tasks–often resulting in *decreased* performance–while the enhanced network visualization designs provided significant performance benefits for both of these sensemaking tasks. These findings support the use of a network visualization based display to present the fusion produced data graph (with task-relevant information graphically incorporated into the network visualization display). Failure to integrate task-relevant information into the display is likely to result in a decreased understanding of the represented information and consequent performance decreases.

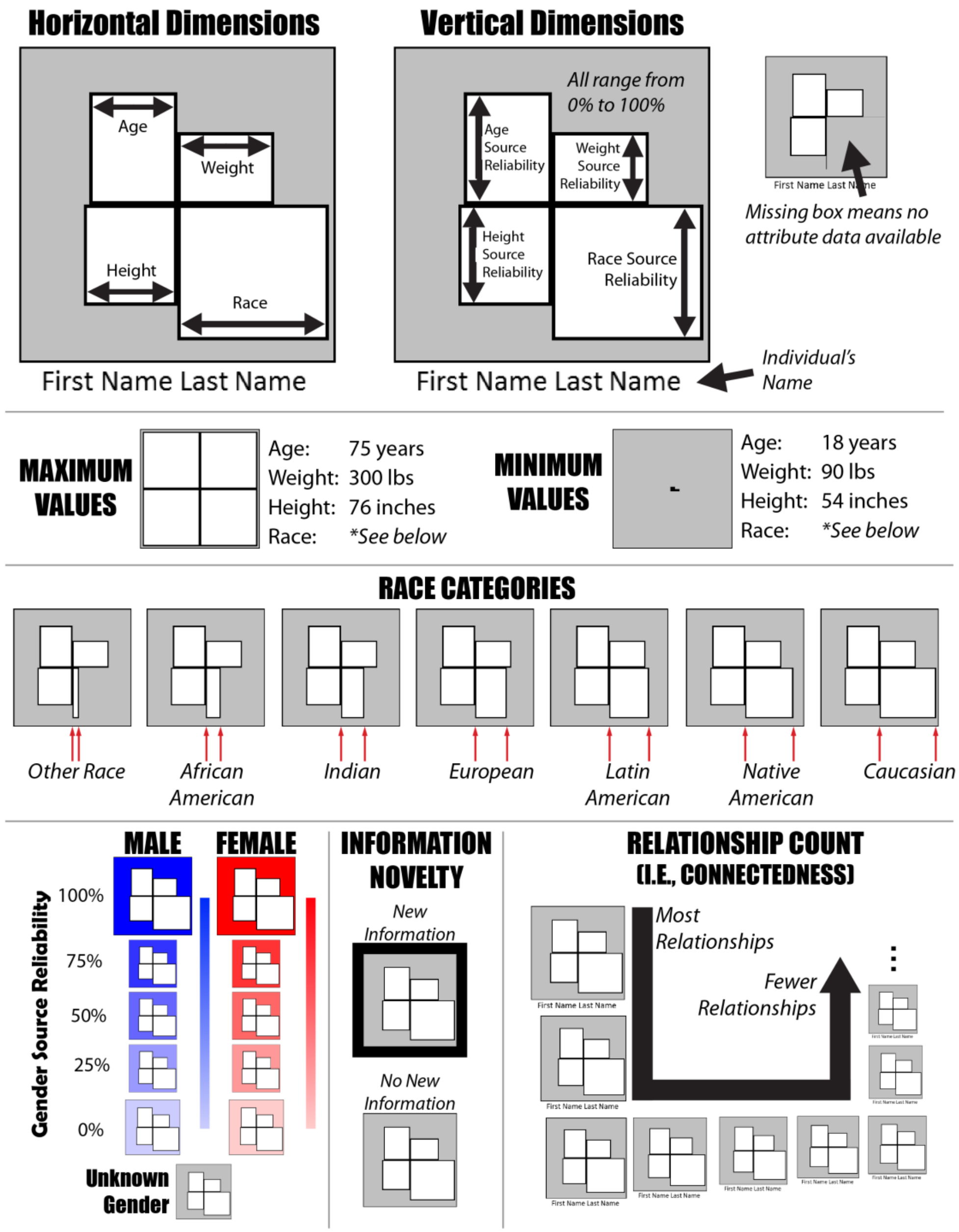


Figure 1: Type 1 network visualization display node graphical dimensions to represented variables mapping legend

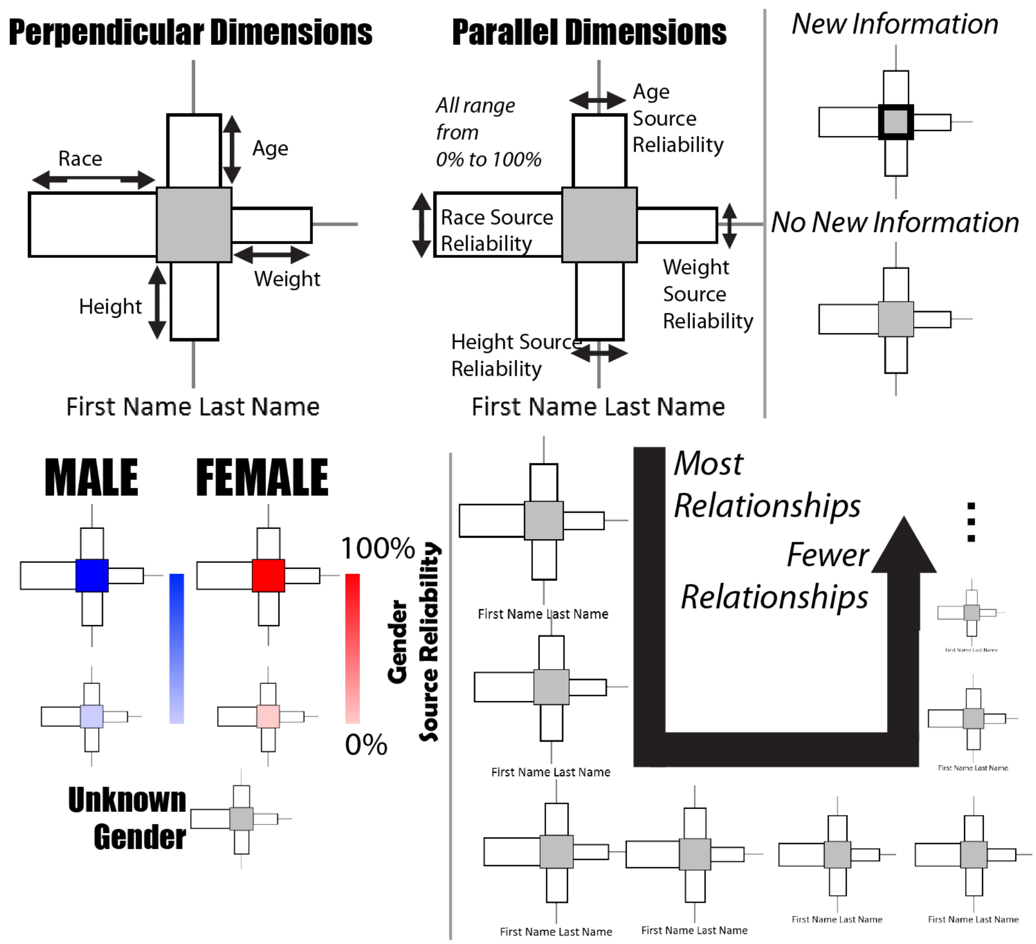


Figure 2: Type 2 network visualization display node graphical dimensions to represented variables mapping legend

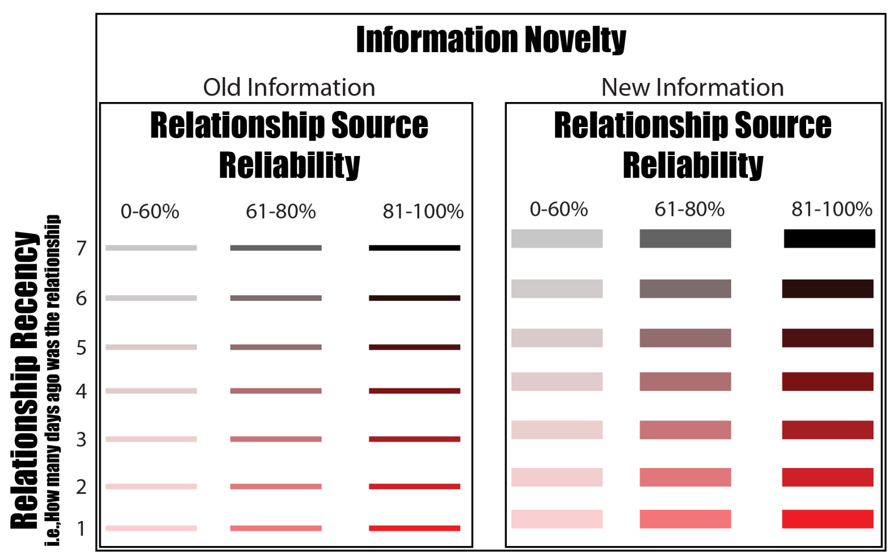


Figure 3: Relationship/link graphic dimension to variable mapping legend

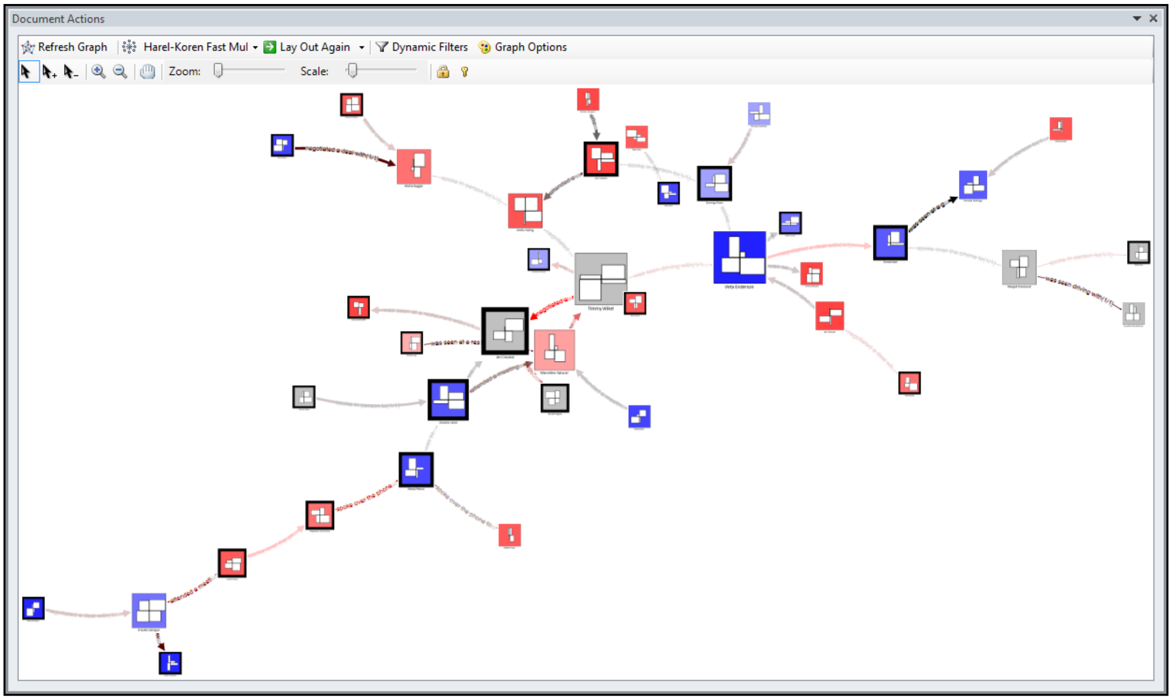


Figure 4: Type 1 network visualization display sample network view

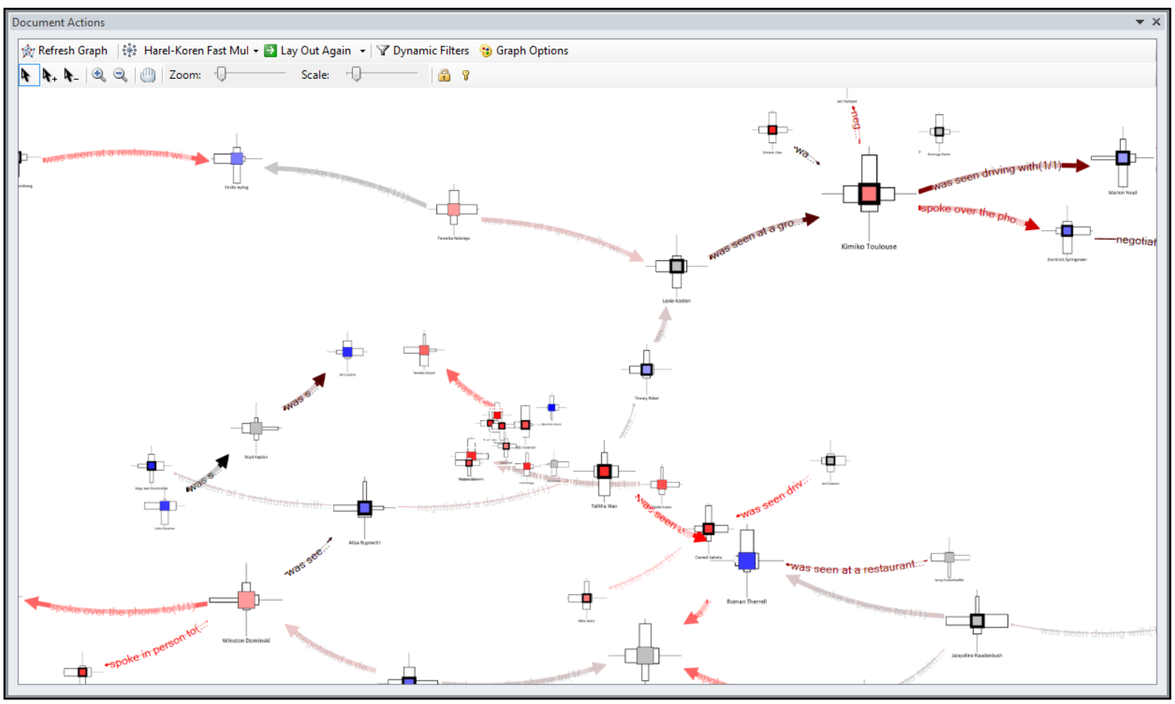


Figure 5: Type 2 network visualization display sample network view

3.1.2 Visual Hard Evidence Extraction and Description

This report describes our progress on visual evidence extraction and description as part of the hard/soft information fusion project. Our work involved two parts: data set study and visual hard evidence extraction and description. Our data set study primarily focused on the DARPA VIRAT

data set for inclusion into the greater project. The visual hard evidence extraction developed and applied state of the art object detectors from computer vision and subsequently generated attributed descriptions of the detected objects. Our report describes this work in detail.

3.1.2.1 Data Set Analysis

The VIRAT dataset was developed in the DARPA VIRAT program. It comprises a large corpus of surveillance videos captured in various environments. The videos are annotated with object bounding boxes (mainly humans and vehicles) as well as action labels (for the VIRAT specific actions). We have analyzed this data set for inclusion into the hard/soft information fusion project. The next two subsections describe our findings of videos that match the SUN messages and those that do not.

3.1.2.1.1 Messages with video matches

We list the videos that have matched messages and include a frame snapshot for each in Figure 6.

1 **Van Parking**

Message: Coalition forces in the Shi'a neighborhood of Abu T'Shir arrest a man after he was observed directing the offload on heavy weapons from a van parked next to a warehouse in Abu T-Shir //MGRSCOORD: 38S MB 43826 78793//

Lat: 33.25712 Long: 44.40640
Day/Time: 15:34:15, (Tuesday) January 26, 2010 (01/26/2010)
Message Number in Complete Set: 3 (SUN3)

Files:

- Video Frame Rate: 30Hz Color
- Mounted Suite: VIRAT_S_050200_00_000106_000380.mp4

2 **Ahmad's vehicle stopped**

Message: GPS tracking device monitoring the movements of Dhanun Ahmad. Mahmud Ahmad detect his vehicle stopped just north of Al-Kut //MGRSCOORD: 38S NB 68893 00519//.

Lat: 32.94877 Long: 45.85308
Day/Time: 13:24:27, (Tuesday) March 16, 2010 (03/16/2010)
Message Number in Complete Set: 248 (SUN39)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_000001.mp4

3 **Ahmad's vehicle stopped**

Message: the GPS tracking Dhanun Ahmad detect his vehicle stopped approximately 60 kilometers east of Al-Kut //MGRSCOORD: 38S NB 83999 55164//.

Lat: 32.94877 Long: 45.85308
Day/Time: 15:35:15, (Tuesday) March 16, 2010 (03/16/2010)
Message Number in Complete Set: 251 (SUN43)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_050201_01_000147_000321.mp4

4 **Ahmad's vehicle stopped**
Message: Informant Dhanun Ahmad leaves a message in his handler's voice drop-box stating his truck is broken down somewhere to the southeast of Badrah. The person that accompanied him has left with their escorts to find a fan belt and water. Another man was left behind keeping armed watch as the area has many bandits; he said he hopes his efforts are appreciated.

Lat: 32.94877 Long: 45.85308
Day/Time: 15:36:15, (Tuesday) March 16, 2010 (03/16/2010)
Message Number in Complete Set: 252 (SUN44)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_050201_02_000395_000483.mp4

5 **Ahmad on the move**
Message: GPS tracker shows Dhanun Ahmad's vehicle on the move heading west approximately 50 kilometers north-east of al-Kut //MGRSCOORD: 38S NB 79738 45931// traveling at 15 km/hr.

Lat: 32.94877 Long: 45.85308
Day/Time: 10:15:16, (Tuesday) March 16, 2010 (03/16/2010)
Message Number in Complete Set: 253 (SUN45)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_000200_01_000226_000268.mp4

6 **Ahmad's vehicle stopped**
Message: GPS detects Ahmad's vehicle stopped approximately 10 km north of Al-Kut.

Lat: 32.94877 Long: 45.85308
Day/Time: 10:17:16, (Wednesday) March 17, 2010 (03/17/2010)

Message Number in Complete Set: 254 (SUN46)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_000200_03_000657_000899.mp4

7 Ahmad on the move

Message: GPS indicates Dhaun Ahmad on the move approximately 12 km north of Al-Kut traveling at 35km/h.

Lat: 32.94877 Long: 45.85308
Day/Time: 10:19:16, (Wednesday) March 17, 2010 (03/17/2010)
Message Number in Complete Set: 256 (SUN48)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_000200_06_001693_001824.mp4

8 Ahmad's vehicle stopped

Message: Vehicle carrying Dhanun Ahmad is stopped approximately 40 km southeast of Baghdad in the village of Dura'iya just south of Salman Pak.

Lat: 32.94877 Long: 45.85308
Day/Time: 10:14:16, (Wednesday) March 17, 2010 (03/17/2010)
Message Number in Complete Set: 257 (SUN49)

Files:

- Video Frame Rate: 30Hz
- Mounted Suite: VIRAT_S_000200_00_000100_000171.mp4

Figure 6: Example snapshots from message matched videos

3.1.2.1.2 Messages with no video match

Many videos in the VIRAT dataset are in parking lots scenarios. Therefore, they do not match with the soft messages from the SUN set. Examples are listed below.

1. (SUN1) 01/25/10 - They have been known to cross sectarian boundaries when they can turn a profit. They have both Shia and Assyrians on the payroll.
2. (SUN5) 01/27/10 - Ahmad Mahmud was placed in custody after his arrest along the Doura Expressway.
3. (SUN6) 01/28/10 - Shia militia member Abdul Jabar, arrested by BCT forces in the Shia neighborhood of Abu TShir.
4. (SUN25) 03/04/10 - The Iraqi police in Karkh report the escape of Dhanun Ahmad Mahmud Ahmad, from police headquarters in Karkh. Iraqi police report Ahmad and three other criminals were freed during an early morning raid and gun-battle.
5. (SUN26) 03/04/10 - Analysis of jail break by Dhanun Ahmad indicates the use of a VBIED detonated at the rear of the police building, followed by a grenade attack to the front of the building. Eye witnesses claim the breakout was conducted by members of a Rashid criminal group. The detainment area was all but destroyed in the attack and several prisoners were killed.
6. (SUN60) 03/24/10 - RT: 2030hrs BCT responding to report of skirmish at the Sunni market in Dora //MGRSCOORD: 38S MB 4362 7988// take control of a truckload of crude chemical weapons.
7. (SUN79) 04/06/10 - ET: 1300hrs BCT forces monitoring safe house in Dora //MGRSCOORD: 38S 43952 80164// report the arrival of six males between the ages of 18-35 in a white minibus with no license plate. The men were photographed as they entered the house.

8. (SUN84) 04/08/10 - BCT analysts conclude cell phone call by Lufti Dilawar on 04/10/10 originated from the BCT monitored safe house in Dora //MGRSCOORD: 38S MB 43952 80164//.
9. (SUN87) 04/11/10 - BCT analysts commence monitoring of all calls in/out of joint ING/INP divisional headquarters in Karkh, Baghdad.
10. (SUN113) 05/02/10 - Review of Iraqi National Guard records indicate ING in Karkh took control of four prisoners from British unit on 04/18/10, however all four were released after database search failed to match detainee bio-data with any known insurgent or criminal activity.

3.1.2.2 Visual Hard Evidence Extraction and Description

3.1.2.2.1 Detection Algorithm and Dataset Description

We use Deformable Part Model, abbreviated DPM, to detect objects in the video frames. The DPM method is the state of the art object detection method in the computer vision literature; it depends heavily on methods for discriminative training and combines a margin-sensitive approach for data mining hard negative examples within a formalism called latent SVM. The DPM model represents an object as a set of parts that are permitted to locally deform allowing it to adapt to variations in object structure, articulations, and weak visual evidence. The model uses histograms of oriented gradients as local features extracted from the images. During inference, the parts are allowed to deform locally and the reported detection score is the one that yields a maximum score over all configurations of the local parts.

We have worked with all three data sets in the project: VIRAT, PSU and TSU datasets. There are 3 scenarios in VIRAT, 1 for a house and parking lot (VIRAT_S_000001 to VIRAT_S_000006), 2 for parking lots (VIRAT_S_000200_00_000100_000171 to VIRAT_S_000207_05_001125_001193, and VIRAT_S_050200_00_000106_000380 to VIRAT_S_050203_09_001960_002083). PSU dataset has 5 scenarios, Arrest At Market_Take#1, Checking Prisoner In Take#3, Jail Break Take#3, Walk Up Deal Take#1_Scene3.1, Walk Up Deal Take#1_Scene3.2. There are 5 scenarios in TSU dataset, Group_activity, Heavy Box Pick Up, Loading & Unloading, Packages Pick Up, Vehicle_flee_1. Figure 7, Figure 8, and Figure 9 depict example detections from each of the three data sets, VIRAT, TSU, and PSU respectively.

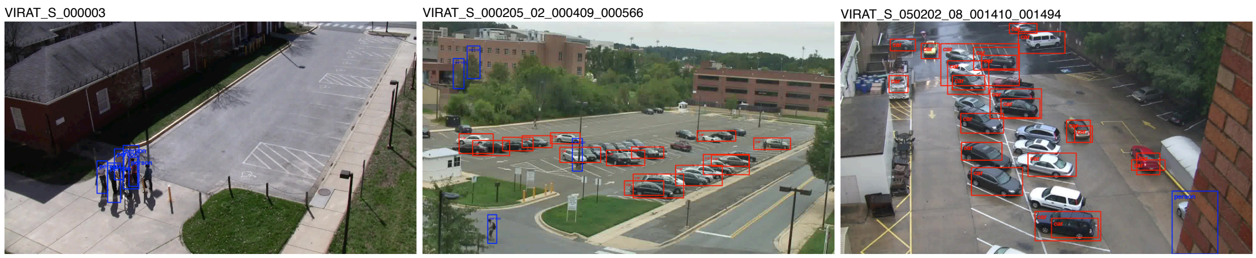


Figure 7: Example detections on the VIRAT videos.

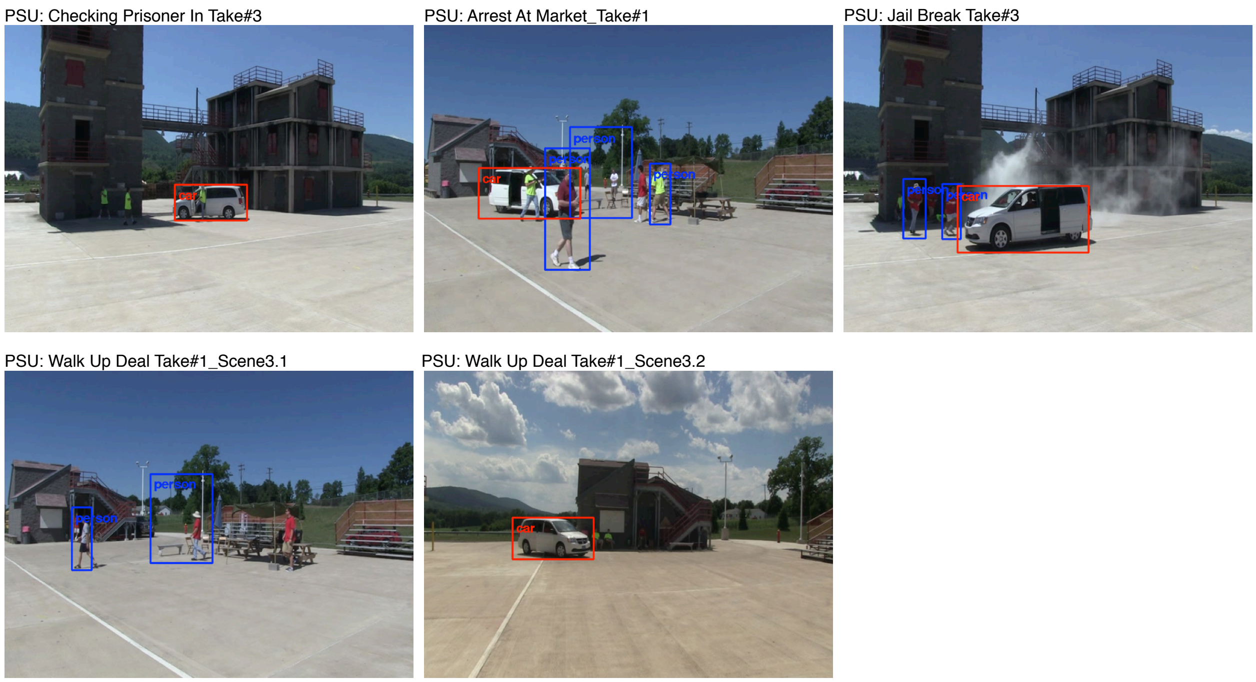


Figure 8: Example detections on the PSU videos.

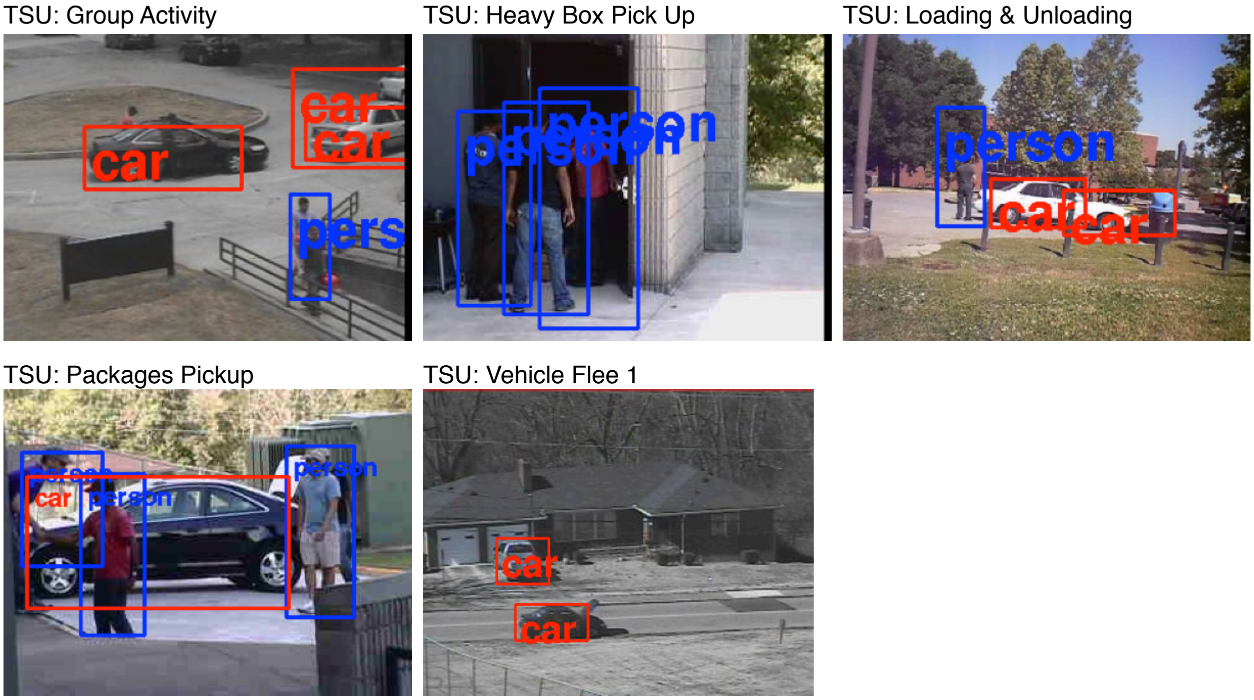
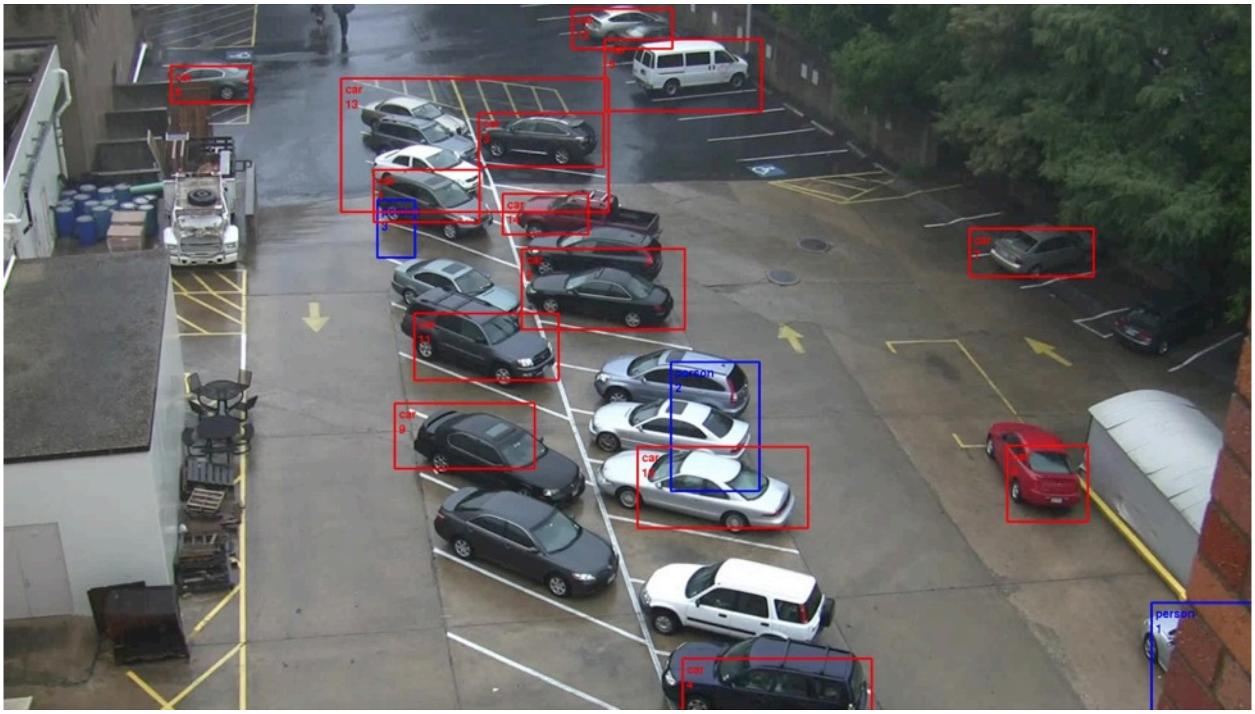


Figure 9: Example detections on the TSU videos.

3.1.2.2.2 Attribute TML file explanation

We have also developed an output mechanism to interface with the hard/soft fusion modules. Our mechanism computes attributes on detected objects and outputs a TML file.

For each video, we detect first detect and track objects. Then, we assign each object a unique ID and compute certain attributes for it. For a person, we use the pseudo-height and for a vehicle, we use the color. Last, we record the hard evidence in a TML file. The following one is a specific example showing the tracking result and TML file content. Figure 10 shows an example video frame, detections, and the generated TML file.



<tml>
<data ref="CAM_UB">1,2010-03-20T13:28.17,3,1,83.36,121.28,141.36,170.28,32.546,47.370,black/deep,32.507,47.285</data>
<data ref="CAM_UB">2,2010-03-20T13:28.18,3,1,83.36,121.28,141.36,170.28,32.546,47.370,black/deep,32.507,47.285</data>
<data ref="CAM_UB">3,2010-03-20T13:28.19,3,1,83.36,121.28,141.36,170.28,32.546,47.370,black/deep,32.507,47.285</data>
<data ref="CAM_UB">4,2010-03-20T13:28.20,3,1,83.36,121.28,141.36,170.28,32.546,47.370,black/deep,32.507,47.285</data>
……

</tml>

Figure 10: Example TML output

Each data tag indicates one object, the reference part, "CAM_UB", represents from the UB team. The first column is frame ID, like 1, 2, 3, 4 in the above example. The second column is time, like 2010-03-20T 13:28.17. Then 9 elements are one unit, the first one is object class ID (1 for person, 2 for bus, 3 for car), then for unique object ID (car 1, car 2, person 1, person2), the following two are the object location, like 83.36, 141.36. Middle four ones are bounding box information, for instance, 141.36, 170.28, 32.546, 47.370. The last one in these 9 elements is the attribute, height (tall/medium/short) for person, color (black/deep, white, red) for vehicles (bus/car). The final two columns are the location for the sensor, like 32.507, 47.285.

1. Person Height: There is no depth information for these video dataset. So we have established a pseudo-height extraction based on the detected object bounding box. In brief, the height of the bounding box is used to determine the outputted height: we

quantize the height based on the viewing frustum and have a lingual determinant associated with each quantized value.

2. Vehicle Color: We compute the RGB color histogram within the bounding box of the detected vehicles. To find the attribute color, we select the maximum mode of the distribution and associate the nearest lingual descriptor for it, such as black, white, red, and so on.
3. Time: The initial time value, such as 2010-03-16 13:23:16, is assigned to the video according to the status when it matches with the SUN soft message. It is then increased with the frames going with the time and record the exact time in some certain format for each specific object in every frame, for example, 2010-03-20T13:28.17.
4. Space: We assign an initial space location value, such as 32.507, 47.285, to the video frame center (the physical central point of all pixels) according to the status when it matches with the SUN soft message. Then, we obtain the relative position with the frame center based on the bounding box information of object. Then, we calculate the exact space coordinate, latitude and longitude as 32.546 and 47.370.

3.1.3 Tractor: Extracting Semantic Information from Soft Data

3.1.3.1 Introduction

Tractor is a system for soft message understanding within the context of "Hard and Soft Information Fusion". Information obtained from physical sensors such as RADAR, SONAR, and LIDAR are considered hard information. Information from humans expressed in natural language is considered soft information. Tractor [C.1.1] is a computational system that understands isolated English intelligence messages in the counter-insurgency domain for later fusion with each other and with hard information, all to aid intelligence analysts to perform situation assessment. In this context, "understanding" means creating a knowledge base (KB), expressed in a formal knowledge representation (KR) language that captures the information in an English message.

Tractor takes as input a single English message. The ultimate goal is for Tractor to output a KB representing the semantic information in that message. Later systems of the larger project combine these KBs with each other and with hard information. Combining KBs from different messages and different hard sources is done via a process of data association that operates by comparing the attributes of and relations among the entities and events described in each KB. It is therefore important for Tractor to express these attributes and relations as completely and accurately as possible.

Our approach is to use largely off-the-shelf software for processing the message text, to be discussed in Section 3.1.3.1.1. The output of text processing is a hybrid syntactic-semantic representation that is mostly syntactic, but contains some semantic information due to the semantic classifications added by named-entity recognizers. We translate the output of the text processing to the KR language we use. The KR language is discussed in Section 3.1.3.1.2, and the translator is discussed in Section 3.1.3.1.3. This KB is enhanced with relevant ontological and geographical information, discussed in Section 3.1.3.1.4. Finally, hand-crafted syntax-

semantics mapping rules, discussed in Section 3.1.3.1.6, are used to convert the mostly syntactic KB into a mostly semantic KB. This is still a hybrid syntactic-semantic representation, because the mapping rules do not convert all the syntactic information. The results of testing and evaluating the system are presented and discussed in Section 3.1.3.1.7. Work on a new KR system, which will facilitate a concurrent approach to syntax-semantics mapping is discussed in Section 3.1.3.1.8. An overview of the Tractor architecture is shown in Figure 11.

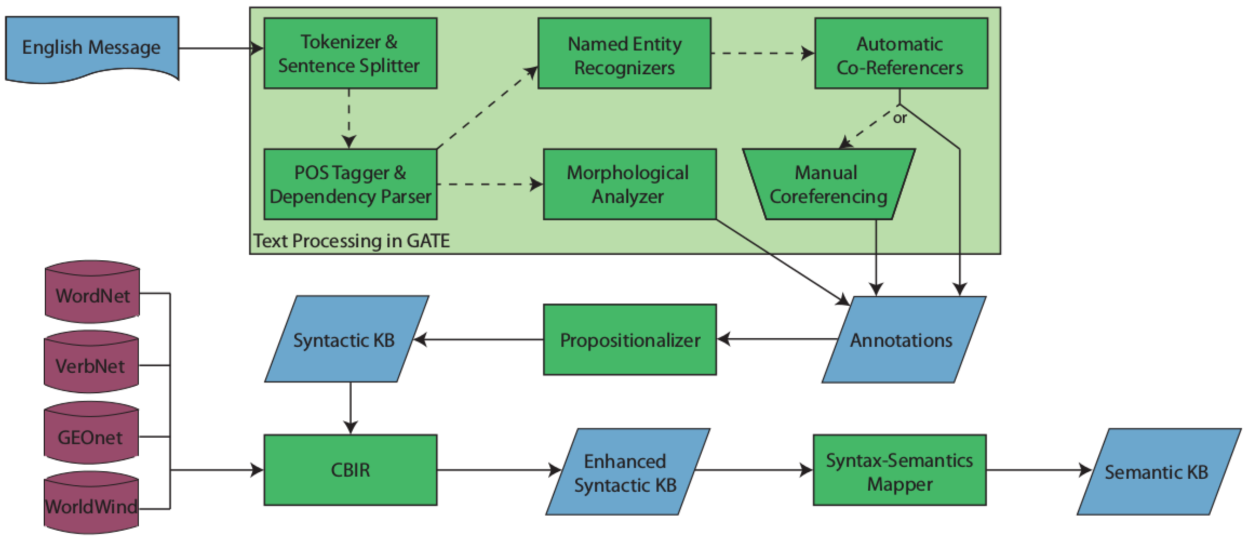


Figure 11: Tractor architecture.

3.1.3.1.1 Text Processing

3.1.3.1.1.1 Introduction

For the text processing phase of Tractor we use GATE, the General Architecture for Text Engineering [C.1.2], which is a framework for plugging in a sequence of "processing resources" (PRs). We currently use eleven such PRs, most of which come from the ANNIE (a Nearly-New Information Extraction System) suite [C.1.3]. The PRs we use include those discussed in the following sections (*cf.* Figure 11).

3.1.3.1.1.2 Tokeniser & Sentence Splitter

The ANNIE English Tokeniser divides the text into words, numbers, and punctuation marks. The ANNIE Sentence Splitter segments the text into sentences. We have modified the ANNIE English Tokeniser to treat times, such as 1:20pm, and time periods such as "1920s" and "mid-20s" as single tokens with features for their components. We have made only minor changes to the sentence splitter.

3.1.3.1.1.3 POS Tagger and Dependency Parser

The Stanford Dependency Parser [C.1.4] identifies the part-of-speech (POS) of each word in the message, and creates a dependency parse of each sentence. We have made no major changes to this PR.

3.1.3.1.1.4 Morphological Analyzer

We tried using the English Snowball Stemmer. However, we found that, even with Paul Bunter's revisions as described in our Year 3 Report, the English Snowball Stemmer was producing results that were not adequate for our purposes. For example, it was giving "overheard" as the stem of "overheard". Therefore, we switched to the GATE Morphological Analyzer. This is a rule-based morphological analyzer, which operates on the tokens produced by the Tokeniser, and uses the part-of-speech tags produced by the Stanford Parser. Since "overheard" is a verb, it produces the correct "overhear" as its root form. We have corrected a few minor bugs in the GATE Morphological Analyzer which prevented it from working on some document types, and added a few rules.

3.1.3.1.1.5 Named-Entity Recognizers

The named-entity recognizers we use are the ANNIE Gazetteer, a list-based named entity recognizer, and the ANNIE NE Transducer which uses a set of JAPE rules to recognize named entities. The changes we made to the named-entity recognizers include additions to the Gazetteer to assist in recognizing car companies and models, additional cities and facilities, roles and names of persons, pronominal references to persons, religious groups, and organization names. We have also added new rules to the NE Transducer to recognize distances (including heights) and weights in standard formats; groups of persons identified by a listing in the message, such as "Dhanun Ahmad Mahmud, Mu'adh Nuri Khalid Jihad, Sattar'Ayyash Majid, Abd al-Karim, and Ghazi Husayn"; names of persons, locations, and organizations with common Arabic prefixes such as "al-"; decade time periods such as "20s"; and dates in context such as "morning of 01/23".

3.1.3.1.1.6 Co-Referencers

The co-referencers we use are: the ANNIE Orthomatcher, which creates co-reference chains (sets of co-referring mentions) for names that are judged to be similar enough; the ANNIE Nominal Coreferencer, which creates co-reference chains for some noun phrases other than names; and the ANNIE Pronominal Coreferencer, which performs anaphora resolution, co-referencing pronouns with each other and with other mentions. We have corrected several errors in the code for these co-referencers which caused processing to halt unexpectedly.

3.1.3.1.1.7 Co-Reference Editor

The co-reference editor provides a GUI within GATE that allows a user to correct and supplement the co-reference chains computed by the three automatic co-referencers discussed in Section 3.1.3.1.1.6. The co-reference editor is optional in that Tractor can be run completely automatically, using only the co-reference PRs discussed in Section 3.1.3.1.1.6, or a user can run all the GATE PRs on one or more messages, then use the co-reference editor on one or more of those messages, then save the results of text processing in XML files, then have Tractor continue the rest of its processing automatically. Tractor can also be run completely automatically, but directed to input and use the stored co-reference chains decided on by an earlier human use of the co-reference editor. The changes we made to the co-reference editor include several bug fixes (some of which have now been incorporated by the developers into the current version of GATE), and user interface enhancements to make it slightly easier to use by being more

consistent with a user's expectations for interface design (namely, popups which would show and hide based on mouse-over timeouts, have been changed to clicks).

3.1.3.1.1.8 Files of Annotations

The final result of GATE is, for each message, an XML file containing a set of annotations, each consisting of an ID number, a Type, a starting and ending position in the sequence of characters of the message, and a set of feature-value pairs. Co-reference chains are recorded by each annotation in a co-reference chain containing a "matches" feature with its value being a list of the IDs of the annotations in the chain. Annotation Types include Person, Location, Organization, JobTitle, Money, Group, and Date, and two more generic Types: Lookup and Token. Additional ontological information is given by some annotations having a feature of majorType, minorType, and/or kind. There is an annotation of type Token for every token recognized by the Tokeniser PR. So, as should be clear, a single span of text might be represented by multiple annotations. Only the start and end positions indicate when an annotation of one PR annotates the same text string as an annotation of another PR.

3.1.3.1.2 SNePS 3

We use SNePS 3 [C.1.5] as the KR system for the KBs created by Tractor from the English messages. SNePS 3 is simultaneously a logic-based, frame-based, and graph-based KR system [C.1.6], and is one of the latest members of the SNePS family of KR systems [C.1.7]. In this report, we will show SNePS 3 expressions using the logical notation, `(R a`$_1$`… a`$_n$`)`, where `R` is an n-ary relation and `a`$_1$`, …, a`$_n$ are its n arguments. We will refer to such an expression as a "proposition". We will use "assertion" to refer to a proposition that is taken to be true in the KB, and say "assert a proposition" to mean adding the proposition to the KB as an assertion. We will also speak of "unasserting a proposition" to mean removing the assertion from the KB. The arguments of a proposition are terms that could denote words, occurrences of words in the message (called "tokens"), syntactic categories, entities in the domain, events in the domain, classes (also referred to as "categories") of these entities and events, or attributes of these entities and events.

We can classify relations, and the propositions in which they occur, as either: **syntactic**, taking as arguments terms denoting words, tokens, and syntactic categories; or as **semantic**, taking as arguments terms denoting entities and events in the domain, their categories, attributes, and properties. A KB is syntactic to the extent that its assertions are syntactic, and is semantic to the extent that its assertions are semantic. The KB first created by Tractor from a message is mostly syntactic. After the syntax-semantics mapping rules have fired, the KB is mostly semantic. A subtle change that occurs as the mapping rules fire is that terms that originally denote syntactic entities are converted into denoting semantic entities.

3.1.3.1.3 The Propositionalizer: Conversion to Propositional Syntactic Graphs

The Propositionalizer examines the annotations produced by the GATE PRs, and produces a set of SNePS 3 assertions. The stages of the Propositionalizer are: annotation merging; correction of minor errors in syntactic categories, particularly when a token is known to be part of a person's name; canonicalization of dates, times, weights, and heights; and processing the structured portion of semistructured messages. Annotations covering the same range of characters are combined into one SNePS 3 token-denoting term. Dates and times are converted

into ISO8601 format. Annotation types, subtypes (where they exist), parts-of-speech, and dependency relations are converted into logical assertions about the tokens. The actual text string of an annotation and the root found by the morphological analyzer are converted into terms and related to the annotation-token by the `TextOf` and `RootOf` relations, respectively. Co-reference chains are converted into instances of the SNePS 3 proposition `(Equiv` $t_1 \ldots t_n$`)`, where $t_1 \ldots t_n$ are the terms for the co-referring tokens. The Propositionalizer supports concurrent processing of GATE XML files through the use of a thread pool.

As an example of the Propositionalizer's output, consider message syn194:

194. 03/03/10 - Dhanun Ahmad has been placed into custody by the Iraqi police and transferred to a holding cell in Karkh; news of his detainment is circulated in his neighborhood of Rashid.

The basic information about the word "placed" in SNePS 3 is

```
(TextOf placed n20)

(RootOf place n20)

(TokenRange n20 38 44)

(SyntacticCategoryOf VBN n20)
```

Here, `n20` is a SNePS 3 term denoting the occurrence of the word "placed" in character positions 38–44 of the message text. The last proposition says that the syntactic category (part of speech) of that token is VBN, the past participle of a verb [C.1.3], Appendix G.

Some of the dependency information about "placed", with the text to make it understandable is

```
(nsubjpass n20 n169)

(TextOf Ahmad n169)

(prep n20 n22)

(TextOf into n22)
```

That is, "Ahmad" is the passive subject of "placed", and "placed" is modified by a prepositional phrase using the preposition "into".

Some of the information about "Karkh" is

```
(TextOf Karkh n182)

(SyntacticCategoryOf NNP n182)

(Isa n182 Location)
```

Notice that in the first two of these assertions, `n182` denotes a token (a word occurrence), but in `(Isa n182 Location)`, it denotes an entity, specifically a location, in the domain. This change in the denotation of individual constants is a necessary outcome of the fact that we form a KB representing the syntactic information in a text, and then gradually, via the syntax-semantics mapping rules, turn the same KB into a semantic representation of the text.

The SNePS 3 KB that results from the Propositionalizer is what we call the syntactic KB. Although it contains some semantic information, such as `(Isa n182 Location)`, most of the information in it is syntactic.

3.1.3.1.4 CBIR Enhancement

Context-Based Information Retrieval (CBIR) [C.1.8], [C.1.9], [C.1.10] enhances the syntactic KB with relevant information of two kinds: ontological taxonomic information is added above the nouns and verbs occurring in the KB; and geographical information is added to geographic place names occurring in the message. The information is "relevant" in the sense that, although CBIR has access to large databases of ontological and geographical information, it adds to the syntactic KB only those data that are connected to the terms already in the syntactic KB. For example, it would add ontological information above the term "truck" only to the KB of a message that mentions a truck, and geographic information about Baghdad only to the KB of a message that mentions Baghdad.

3.1.3.1.4.1 Enhancing with Ontological Information

We first used Cyc as the source of ontological information. In Year 5, we switched to WordNet and VerbNet. A comparison of these two sources of ontological information is in Section 3.1.3.1.8.

When we used Cyc, CBIR looked up each noun and verb in ResearchCyc [http://research.cyc.com/] to find the corresponding Cyc concept(s). Then it added to the KB the terms above those concepts in OpenCyc [http://www.opencyc.org/]. Using WordNet and VerbNet, CBIR first looks up in WordNet [C.1.11] all the common nouns that are in a syntactic KB, and adds to the KB the synsets of the nouns, their hypernyms, the hypernyms of their hypernyms, etc., all the way to the top of the ontology. Then it looks up in VerbNet [C.1.12] all the verbs in the KB, and adds all their classes, parent classes, etc. At the top of the VerbNet hierarchy, CBIR looks up all the member verbs of the highest level classes in WordNet, and adds the connected WordNet hierarchy to the VerbNet hierarchy.

Although VerbNet and WordNet are often viewed as hierarchies of words, and thus in the syntactic realm, WordNet synsets are groups of synonymous words "expressing a distinct concept" [C.1.11] and the hypernym relation is a semantic relation between concepts. VerbNet classes are an extension of Levin classes [C.1.13], which add subclasses to "achieve syntactic and semantic coherence among members of a class" [C.1.11]. Thus, the VerbNet and WordNet hierarchies added by CBIR constitute an ontology in the semantic realm. The addition of this ontology adds to the categorization of entities and events begun by the named-entity recognizers. These categories are used by the syntax-semantics mapping rules so that they apply to classes of entities and events, not just to specific ones. In addition, the ontology is used by the scoring algorithms of the data association routine to assess the semantic distance between entities and events mentioned in different messages.

3.1.3.1.4.2 Enhancing with Geographic Information

CBIR looks up every proper noun that is in the message in the NGA GEOnet Names Server database [C.1.14]. To reduce the confusion caused when one name is the name of multiple places, we use our knowledge of our domain to restrict the database to places in Iraq. The information found is added to the KB for the message. For example, looking up Badrah, CBIR

finds that it is a second order administrative division, its MGRS (Military Grid Reference System) coordinates are 38SNB8399760885, its latitude is 33.08333, and it longitude is 45.90000.

If CBIR finds MGRS coordinates, but no latitude and longitude (This particularly happens when MGRS coordinates are explicitly included in a message.), it converts the MGRS coordinates to latitude and longitude using NASA's World Wind software [C.1.15].

For example, the information added about Karkh is

```
(Isa Karkh SectionOfPopulatedPlace)
(GeoPosition Karkh (GeoCoords 33.3217 44.3938))
(MGRS Karkh 38SMB4358187120)
```

The information added by CBIR is important to the data association task in deciding when terms from different messages should be considered to be co-referential.

3.1.3.1.4.3 Pedigree of Information

CBIR can add to the information it contributes meta-information about where that information came from. For example

```
(source (Isa Karkh SectionOfPopulatedPlace) GeoNet)
```

says that the information that Karkh is a section of a populated place came from GeoNet. Other information CBIR adds is noted as coming from CBIR.

If this meta-information is not desired, it can be turned off by a configuration flag.

3.1.3.1.5 Representation Issues

3.1.3.1.5.1 Major Categories of Entities and Events

The actual message texts determine what categories of entities and events appear in the semantic KBs. For example, in the message, "*Owner of a grocery store on Dhubat Street in Adhamiya said ...*", there is a mention of an entity which is an instance of the category store. So the category of stores is represented in the semantic KB. Nevertheless, there are some categories that play a role in the mapping rules in the sense that there are rules that test whether some term is an instance of one of those categories. Such major categories of entities include: Person; Organization (a subcategory of Group); company; Location; country; province; city; Date; Time; Phone (the category of phone instruments); PhoneNumber (the category of phone numbers); MGRSToken; JobTitle; Dimension (such as age, height, and cardinality); Group (both groups of instances of some category, such as "mosques," and groups of fillers of some role, such as "residents"); ReligiousGroup (such as "Sunni"); and extensionalGroup (a group explicitly listed in a text, such as, "*Dhanun Ahmad Mahmud, Mu'adh Nuri Khalid Jihad, Sattar 'Ayyash Majid, Abd al-Karim, and Ghazi Husayn*.") Major categories of events include: Action (such as "break" and "search"); ActionwithAbsentTheme (such as "denounce" and "report"); actionWithPropositionalTheme (such as "say" and "hear"); Perception (such as "learn" and "recognize"); and Event itself.

3.1.3.1.5.2 Functional Term

We use one functional term: (GeoCoords x y) denotes the geographic coordinate position whose latitude is x and whose longitude is y. This is used as the argument of the GeoPosition attribute, as shown above.

3.1.3.1.5.3 Relations

Relations used in the syntactic and semantic KBs can be categorized as either syntactic relations or semantic relations. The syntactic relations we use include the following.

- (TextOf *x y*) means that the token *y* in the message is an occurrence of the word *x*.
- (RootOf *x y*) means that *x* is the root form of the word associated with token *y*.
- (SyntacticCategoryOf *x y*) means that *x* is the syntactic category (part-of-speech) of the word associated with token *y*.
- (*r x y*), where *r* is one of the dependency relations listed in [C.1.4], for example nsubj, nsubjpass, dobj, prep, and nn, means that token *y* is a dependent of token *x* with dependency relation *r*.

The semantic relations we use include the ones already mentioned (such as Isa and Equiv), and the following.

- (Type *c1 c2*) means that *c1* is a subcategory of *c2*.
- (hasName *e n*) means that *n* is the proper name of the entity *e*.
- (GroupOf *g c*) means that *g* is a group of instances of the class *c*.
- (GroupByRoleOf *g r*) means that *g* is a group of entities that fill the role, *r*.
- (MemberOf *m g*) means that entity *m* is a member of the group *g*.
- (hasPart *w p*) means that *p* is a part of entity *w*.
- (hasLocation *x y*) means that the location of entity *x* is location *y*.
- (Before *t1 t2*) means that time *t1* occurs before time *t2*.
- (*r x y*), where *r* is a relation (including possess, knows, outside, per-country_of_birth, org-country_of_headquarters, agent, experiencer, topic, theme, source, and recipient), means that the entity or event *x* has the relation *r* to the entity or event *y*.
- (*a e v*), where *a* is an attribute (including cardinality, color, Date, height, GeoPosition, sex, per-religion, per-date_of_birth, and per-age), means that the value of the attribute *a* of the entity or event *e* is *v*.

One relation, although syntactic, is retained in the semantic KB for pedigree purposes: (TokenRange *x i j*) means that the token *x* occurred in the text starting at character position *i*, and ending at character position *j*. This is retained in the semantic KBs so that semantic information may be tracked to the section of text which it interprets. Two other syntactic relations, TextOf and RootOf, are retained in the semantic KB at the request of the data association group to provide term labels that they use for comparison purposes.

We believe that the syntactic relations we use are all that we will ever need, unless we change dependency parsers, or the dependency parser we use is upgraded and the upgrade includes new dependency relations. However, we make no similar claim for the semantic relations.

Assertions that use syntactic relations are called "syntactic assertions," and those that use semantic relations are called "semantic assertions."

3.1.3.1.5.4 Representation of Events

To represent events, we use a neo-Davidsonian representation [C.1.17], in which the event is reified and semantic roles are binary relations between the event and the semantic role fillers. For suggestions of semantic roles, we have consulted the entries in the Unified Verb Index [C.1.18]. For example, in the semantic KB Tractor constructed from message syn064,

> 64. 01/27/10 - BCT forces detained a Sunni munitions trafficker after a search of his car netted IED trigger devices. Ahmad Mahmud was placed in custody after his arrest along the Dour'a Expressway, //MGRSCOORD: 38S MB 47959 80868//, in East Dora.

the information about the detain event includes

```
(Isa n18 detain)
(Date n18 20100127)
(agent n18 n16)
(GroupOf n16 force)
(Modifier n16 BCT)
(theme n18 n26)
(Equiv n230 n26)
(Isa n230 Person)
(hasName n230 "Ahmad Mahmud")
```

That is, `n18` denotes a detain event that occurred on 27 January 2010, the agent of which was a group of BCT forces, and the theme of which was (coreferential with) a person named Ahmad Mahmud.

3.1.3.1.5.5 Source Information

It was mentioned in Section 3.1.3.1.4 that the relation `source` is used by CBIR to indicate where information came from. That same relation is used to indicate the source of information contained in the messages. For example, message syn063 contains the sentences, "A man arrested in the Shi'a neighborhood of Abu T'Shir … has been identified as Abdul Jabar. Jabar claims he lives in the neighborhood." Using the `source` relation, Tractor indicates that Abdul Jabar is the source of the information that Abdul Jabar, himself, lives in Abu T'Shir.

3.1.3.1.6 Syntax-Semantics Mapping

The purpose of the syntax-semantics mapping rules is to convert information expressed as sets of syntactic assertions into information expressed as sets of semantic assertions. The rules were hand-crafted by examining syntactic constructions in subsets of our corpus, and then expressing the rules in general enough terms so that each one should apply to other examples as well.

The rules are tried in order, so that earlier rules may make adjustments that allow later rules to be more general, and earlier rules may express exceptions to more general later rules. As of this writing, there are 189 mapping rules, which may be divided into several categories:

- CBIR, supplementary enhancement rules add ontological assertions that aren't found by CBIR, but that relate to terms in the message;
- SYN, syntactic transformation rules examine syntactic assertions, unassert some of them, and make other syntactic assertions;
- SEM, semantic transformation rules examine semantic assertions, unassert some of them, and make other semantic assertions;
- SYNSEM, true syntax-semantic mapping rules examine syntactic assertions and maybe some semantic assertions as well, unassert some of the syntactic assertions, and make new semantic assertions;
- CLEAN, cleanup rules unassert some remaining syntactic assertions that do not further contribute to the understanding of the message;
- INFER, inference rules make semantic assertions that are implied by other semantic assertions in the KB.

Due to space constraints, only a few rules will be discussed. An example of a syntactic transformation rule is

```
(defrule passiveToActive
   (nsubjpass ?verb ?passsubj)
   =>
   (assert ‘(dobj ,?verb ,?passsubj))
   (unassert ‘(nsubjpass ,?verb ,?passsubj))
   (:subrule
        (prep ?verb ?bytok)
        (TextOf by ?bytok)
        (pobj ?bytok ?subj)
   =>
        (assert ‘(nsubj ,?verb ,?subj))
        (unassert ‘(prep ,?verb ,?bytok))
```

```
(unassert ‘(pobj ,?bytok ,?subj))))
```

This rule would transform the parse of “*BCT is approached by a man*” to the parse of “*a man approached BCT*”. The rule fires even if the “by” prepositional phrase is omitted.

There are also some rules for distribution over conjunctions. One such rule would transform the parse of “*They noticed a black SUV and a red car parked near the courthouse*” to the parse of “*They noticed a black SUV parked near the courthouse and a red car parked near the courthouse*” by adding an additional `partmod` relation, from the token for “car” to the head token of “*parked near the courthou*se”. Then another rule would transform that into the parse of “*They noticed a black SUV parked near the courthouse and they noticed a red car parked near the courthouse*” by adding a second `dobj` relation, this one from the token of “*noticed*” to the token of “*car*.”

Some examples of true syntax-semantics mapping rules operating on noun phrases (presented in the relative order in which they are tried) are:

```
(defrule synsemReligiousGroup
   (Isa ?g relig_group_adj)
   (TextOf ?name ?g)
   =>
   (assert ‘(Isa ,?g ReligiousGroup))
   (assert ‘(hasName ,?g ,?name))
   (assert ’(Type ReligiousGroup Group))
   (unassert ‘(Isa ,?g relig_group_adj)))
```

This rule would transform the token for “*Sunni*”, which the GATE named entity recognizers recognized to name a relig_group_adj, into an entity that is an instance of `ReligiousGroup`, whose name is `Sunni`. It also makes sure that the relevant fact that `ReligiousGroup` is a subcategory of `Group` is included in the semantic KB for the current message.

```
(defrule hasReligion
   (Isa ?religiongrp ReligiousGroup)
   (nn ?per ?religiongrp)
   (hasName ?religiongrp ?religion)
   =>
   (assert ‘ (MemberOf ,?per ,?religiongrp))
   (assert ‘ (per-religion ,?per ,?religion))
   (unassert ‘ (nn ,?per ,?religiongrp)))
```

This rule would assert about the token of "*youth*" in the parse of "*a Sunni youth*" that it is a member of the group named `Sunni`, and that its religion is `Sunni`. It also would unassert the nn dependency of the token of "*Sunni*" on the token of "*youth*".

```
(defrule properNounToName
   (SyntacticCategoryOf NNP ?token)
   (TextOf ?text ?token)
   =>
   (assert ‘(hasName ,?token ,?text))
   (unassert ‘(SyntacticCategoryOf NNP ,?token))
   (unassert ‘(TextOf ,?text ,?token)))
```

This rule would transform a token of the proper noun "*Khalid Sattar*" into a token denoting the entity whose name is "`Khalid Sattar`".

```
(defrule nounPhraseToInstance
   (SyntacticCategoryOf NN ?nn)
   (:when (isNPhead ?nn))
   (RootOf ?root ?nn)
   (:unless (numberTermp ?root))
   =>
   (assert ‘(Isa ,?nn ,?root))
   (unassert ‘(SyntacticCategoryOf NN ,?nn))
   (unassert ‘(RootOf ,?root ,?nn)))
```

This rule would transform the token of "*youth*" in the parse of "*a Sunni youth*" into an instance of the category `youth`. The function `isNPhead` returns True if its argument is the head of a noun phrase, recognized by either having a det dependency relation to some token, or by being an nsubj, dobj, pobj, iobj, nsubjpass, xsubj, or agent dependent of some token. (In the corpus we work on, determiners are sometimes omitted.) The `(:unless (numberTermp ?root))` clause prevents a token of a number from being turned into an instance of that number.

Another rule makes the token of a verb an instance of the event category expressed by the root form of the verb. For example, a token of the verb "*detained*" would become an instance of the event category `detain`, which is a subcategory of `Action`, which is a subcategory of `Event`.

Some examples of syntax-semantics mapping rules that analyze clauses (presented in the relative order in which they are tried) are:

```
(defrule subjAction
```

```
   (nsubj ?action ?subj)
   (Isa ?action Action)
   =>
   (assert ‘(agent ,?action ,?subj))
   (unassert ‘(nsubj ,?action ,?subj)))
```

This rule would make the subject of "*detained*" the agent of a `detain` Action-event.

```
(defrule subjPerception
   (nsubj ?perception ?subj)
   (Isa ?perception Perception)
   =>
   (assert ‘(experiencer ,?perception ,?subj))
   (unassert ‘(nsubj ,?perception ,?subj)))
```

This rule would make the subject of "*overheard*" the experiencer of a `overhear` Perception-event.

Another rule makes the date of an event either the date mentioned in the dependency parse tree below the event token, for example the date of the capture event in "*Dhanun Ahmad Mahmud Ahmad, captured on 01/27/10, was turned over to ...*" is `20100127`, or else the date of the message being analyzed.

A final set of syntax-semantics mapping rules convert remaining syntactic assertions into "generic" semantic assertions. For example, any remaining prepositional phrases, after those that were analyzed as indicating the location of an entity or event, the "*by*" prepositional phrases of passive sentences, etc., are transformed into semantic assertions using the preposition as a relation holding between the entity or event that the PP was attached to and the object of the preposition.

As syntax-semantics mapping rules convert syntactic information into semantic information, semantic transformation rules move some of that information to their proper places. One example is

```
(defrule carModelHead
   (Isa ?c CarModel)
   (:when (isNPhead ?c))
   (TextOf ?m ?c)
   =>
   (assert ‘ (Isa ,?c vehicle))
```

```
(assert ‘ (model ,?c ,?m))
(unassert ‘ (Isa ,?c CarModel))
(unassert ‘ (TextOf ,?m ,?c)))
```

This rule applies when the head of a noun phrase is a car model, such as “the 1998 Toyota Corolla driven by Dhanun Ahmad.” The rule `nounPhraseToInstance` would have interpreted this phrase as referring to an instance of Corolla. This rule corrects that to be an instance of vehicle that has Corolla as its model.

Cleanup rules unassert syntactic assertions that were already converted into semantic assertions, for example unasserting `(TextOf x y)` and `(RootOf x y)` when `(Isa y x)` has been asserted. Other cleanup rules unassert remaining syntactic assertions that do not contribute to the semantic KB, such as the `SyntacticCategoryOf` assertions.

The inference rules make certain derivable assertions explicit for the benefit of the data association operation. For example, the agent of an event that occurred at some location on some date was at that location on that date, and the member of a group g_1 that is a subgroup of a group g_2 is also a member of g_2.

3.1.3.1.6.1 Use of Background Knowledge in Syntax-Semantics Mapping

Graded, descriptive adjectives provide linguistic values for attributes of instances of categories such that the adjective and the category imply the attribute [C.1.19], pp. 48ff. The mapping rules use a database of *adjective* × *category* → *attribute* mappings to find the correct attribute. For example, “a young man” is interpreted as a man whose age is young, and “a large gathering” is interpreted as a group whose cardinality is large.

Sometimes a possessive construction indicates ownership, sometimes the part-of relation, and sometimes a weaker association. The mapping rules use a mereological database of parts and wholes to interpret, for example, the phrase “the man’s arm” as an arm that is part of the man, rather than an arm that is owned by the man.

Count nouns (like “car”) denote categories whose instances occur in discrete units that can be counted. Mass nouns denote substances (like “wood) that objects may be made of. Some nouns can be used both ways (“a piece of a cake” vs. “a piece of cake”). The mapping rules use a list of mass nouns, so that, for example, “a man with dark hair” is correctly interpreted as a man who has as a part something which is made of hair whose color is dark. (Notice that this interpretation also makes use of the mereology and the database of graded, descriptive adjectives.)

A common noun, especially one that is the head of a noun phrase, usually denotes an entity that is an instance of the category expressed by the noun. However, the named-entity recognizers recognize certain nouns as job titles, and in that case, the mapping rules identify the noun as denoting an entity that fills the role. For example, in the sentence, “The assistant said the man she treated was covered in dust,” “the man” is understood to denote an instance of the category man, but “the assistant” is understood to denote a person who fills the role of assistant. Similarly, a plural common noun, such as “heavy weapons” is understood to denote a group whose members are instances of the category expressed by the noun (“weapon”), but a plural job title,

such as "BCT analysts" is understood to denote a group whose members fill the role expressed by the job title ("analyst").

Noun phrases that name vehicles have their own peculiar structure that can include color, model year, make, model, and body style. For example, all are included in the phrase "his black 2010 Ford Escape SUV." The named-entity recognizers within GATE recognize colors, years, car companies, car models, and car body styles, and a special mapping rule relates each appropriately to the named entity. If a movement event is modified by a movement preposition whose object is a location, then the location is understood to form the path of the movement. For example, in the sentence, "Dillinger was last seen driving his black 2010 Ford Escape SUV westward down Indianapolis Road at 1:20pm on 3/17/2013," Tractor understands that Indianapolis Road forms the path of the driving event. Moreover, because "westward" is a direction and an adverbial modifier of "driving," the direction along the path is understood to be westward.

If a search of some place uncovers some object, then the object was located in the searched place. For example, Tractor infers from "a search of his car netted IED devices" that the IED devices were located in the car.

Understanding Noun-Noun Modification

The modification of a noun by another noun is used to express a wide variety of semantic relations. However, certain cases are recognized by the mapping rules from the categories of the nouns. (Though exceptions might still occur.) If both nouns denote locations, then the location of the modifying noun is located within the location of the head noun. For example in "Rashid, Baghdad," Rashid is understood as a neighborhood within Baghdad.

However, buildings and other facilities are also locations. (One can be in or next to a building.) So if the head noun denotes a facility, then the facility is understood as being in the location of the modifying noun. For example, "Second District Courthouse" is interpreted as a courthouse located in the Second District.

If the modifying noun is a location, but the head noun is not, then the entity denoted by the head noun is understood as headquartered in the location expressed by the modifying noun. For example "A Baghdad company" is interpreted as a company headquartered in Baghdad.

If neither noun is a location, but both are proper nouns, then they are both assumed to be names of the denoted entity. For example, "Ahmad Mahmud" is interpreted as a person who has both "Ahmad" and "Mahmud" as names, as well as having the full name "Ahmad Mahmud."

If the head noun denotes a person and the modifying noun denotes the name of a religious group (recognized by the named-entity recognizers), then a mapping rule asserts that the person is a member of the religious group and has that religion. For example, "a Sunni munitions trafficker" is understood to be a munitions trafficker whose religion is Sunni and who is a member of the religious group whose name is "Sunni."

If both nouns denote groups, then the head noun is understood to denote a group that is a subgroup of the group denoted by the modifying noun. For example, "BCT analysts" is interpreted to denote a group of analysts all of whom are members of the organization named "BCT".

If the modifying noun denotes an organization, but the head noun does not, then the entity denoted by the head noun is understood to be a member of the organization. For example “the ISG affiliate” is interpreted to be someone filling the role of affiliate within the organization named “ISG.”

Understanding Copulas

If there is a copula between a subject and a noun, then the subject is understood to be co-referential with an entity that is an instance of the category that the predicate noun denotes. For example, “the rented vehicle is a white van” is interpreted to mean that one entity is both a rented vehicle and a white van.

Typically, one end of a dimension has a linguistic value that can be used in neutral questions to ask what value some entity has on that dimension. For example, “How old is he?” is a neutral question about the person’s value on the age dimension without implying that the person is old, but “How young is he?” also suggests that the person is young. Similarly, “How tall is she?” is a neutral question, whereas “How short is she?” suggests that she is short. The neutral value can be used in a copula to say that the subject entity has that value on the implied scale, for example, “Dillinger is old,” is interpreted to mean that Dillinger’s age has the linguistic value “old,” but can also be modified by a specific value to indicate the value on the implied scale. For example, “Dillinger is 30 years old” is interpreted to mean that Dillinger’s value on the age attribute is 30 years. One wouldn’t normally say something like “Dillinger is 20 years young,” or “Betty is 5 feet short.”

Predicate adjectives that do not imply a specific attribute dimension are interpreted as simple properties of the subject. For example “he is secretive” is interpreted to mean that he has the property “secretive”, and “he is apolitical” is interpreted to mean that he has the property “apolitical.”

Making Inferences

The ontology includes a category of symmetric relations so that a particular representational scheme can be used for them [C.1.20]. For example, “match” is symmetric, so the relation expressed in the sentence “The trigger devices netted in the arrest of Dhanun Ahmad Mahmud Ahmad on 01/27/10 match materials found in the truck of arrested ISG affiliate Abdul Wahied.” is represented in such a way that both “the devices match the materials” and “the materials match the devices” are represented. (That is, they match each other.)

Concrete participants in an act performed at some location were at that location at the time of the act. For example, “Ahmad Mahmud was arrested at Expressway on 20100127” is understood to imply that Ahmad was located at the Expressway on 20100127.

If someone drives a vehicle at some time, then the vehicle is not only the object of the driving, it is also the location of the driver at that time. For example, in the sentence, “Dillinger was last seen driving his black 2010 Ford Escape SUV westward down Indianapolis Road at 1:20pm on 3/17/2013,” Dillinger is understood both to be the driver of the SUV and to be located in the SUV at 1320 on 20130317.

The location relation is transitive. So, when interpreting the above sentence, Tractor understands that Dillinger was on the Indianapolis Road at 1320 on 20130317. The subgroup

relation is also transitive. So if Ahmad is a member of one group that is a subgroup of another group, then Ahmad is a member of both groups.

3.1.3.1.7 The Description Facility

The describe AllInstances function displays several lines of text for every set of co-referring terms in the semantic KB such that: they are an instance of some category; and each is associated with its token range. For each set of co-referring terms the following lines are printed:

- Its “best name”
- Its “best class(es)”
- A set of lines describing
    - Its attributes
    - The relations it participates in
- One line for each co-referring term, showing
    - The actual term
    - The token range
    - The message text in that range ordered by the beginning of the token range.

The best class(es) of a set of co-referring terms is the set of the least general classes they are instances of. That is, no class is included in the set if it is a superclass of another class in the set.

The best name of a set of co-referring terms is computed as follows:

- If any of the terms, m, is in the relation (hasName m name), then the best name is the longest of those names.
- If the lexicographically least term name is not of the form n*x*, where *x* is some integer, use that name.
- If the lexicographically least term name is of the form n*x*, where *x* is some integer, choose one of the set of best classes, c, and use the concatenation c*x* as the best name.

The attributes of, and relations involving a term are described by generating a sentence from the docstring of the caseframe of the assertion stating that attribute or relation. Recall that each docstring is a clause with indications of where the fillers of a slot are to be placed. For describeAllInstances, that place is filled by the best name of the filler.

Following are some examples of the output of describeAllInstances for syn194, discussed above:

```
People
==============
|Dhanun Ahmad|
Instance of: Person
place20 has the relation theme to |Dhanun Ahmad|.
transfer37 has the relation theme to |Dhanun Ahmad|.
|Dhanun Ahmad| has the relation possess to Rashid.
```

|Dhanun Ahmad| has the relation possess to detainment58.

|Dhanun Ahmad|'s sex is male.

Coreference Chain:

n196 16-28 "Dhanun Ahmad"

n179 23-28 "Ahmad"

n188 131-134 "his"

n189 169-172 "his"

Locations

==============

Karkh

Instance of: SectionOfPopulatedPlace

Karkh's MGRS is 38SMB4358187120.

Karkh's GeoPosition is latitude |33.32174|, longitude |44.39384|.

cell45 is located at Karkh.

Coreference Chain:

n197 116-121 "Karkh"

Things

==============

cell45

Instance of: cell

transfer37 has the relation recipient to cell45.

cell45 is located at Karkh.

Coreference Chain:

n45 108-112 "cell"

Events

==============

transfer37

Instance of: transfer

transfer37 has the relation recipient to cell45.

transfer37's Time is time1422.

transfer37's Date is |20100303|.

transfer37 has the relation theme to |Dhanun Ahmad|.

transfer37 has the relation agent to police32.

Coreference Chain:

n37 83-94 "transferred"

3.1.3.1.8 Evaluation

3.1.3.1.8.1 Effectiveness of Syntax-Semantics Mapping

The mapping rules were developed by testing Tractor on several corpora of messages, examining the resulting semantic KBs, finding cases where we were not happy with the results, examining the initial syntactic KBs, and modifying or adding to the rule set so that an acceptable result was obtained. These "training" messages included: the 100 messages from the Soft Target Exploitation and Fusion (STEF) project [C.1.21]; the 7 Bomber Buster Scenario messages [C.1.24]; 13 messages of the Bio-Weapons Thread, 84 messages of the Rashid IED Cell Thread, and 114 messages of the Sunni Criminal Thread (SUN), of the 595-message SYNCOIN dataset [C.1.22], [C.1.23]. None of these messages were actual intelligence messages, but are "a creative representation of military reports, observations and assessments" [C.1.23].

In this section, we present an evaluation of how general the mapping rules are, and whether they are perhaps overly general. The generality of the rules were tested through examination of how often the mapping rules fire on a "test" dataset not previously examined. We'll look at the amount of syntactic and semantic data there are in the processed graphs from our test and training sets. We'll also look at how many mistakes Tractor makes on the test dataset, to test for over-generality. Combined, these three experiments show that our rules are general, but not overly so, that the amount of semantic data in the resultant semantic KBs is quite high, and that the degree of semantization compares well with that of our training sets.

We begin by addressing the question of, given that the mapping rules were developed using the training messages, how general are they? To what extent do they apply to new, unexamined, "test" messages? To answer this question, we used the 57 messages of the Sectarian Conflict Thread (SCT) of the SynCOIN dataset. These messages, averaging 46 words per message, contain human intelligence reports, "collected" over a period of about five months, which describe a conflict among Christian, Sunni, and Shi'a groups. The messages describe events in detail, and entities usually only through their connection to some group or location.

Table 1: The number of mapping rules in each category, the number of those rules that fired on any message in the SCT dataset, the total number of times those rules fired, and the average number of times they fired per message.

Rule Type	Rule Count	Rules Fired	Times Fired	Firings/message
CBIR	1	1	474	8.32
SYN	23	13	1,596	28.00
SEM	5	5	328	5.75

SYNSEM	99	56	2,904	50.95
INFER	9	8	135	2.37
CLEAN	10	8	6,492	113.89
TOTAL	**147**	**91**	**11,929**	**209.28**

We divided the rules into the six categories listed in Section 3.1.3.1.6, and counted the number of rules used in the SCT corpus, along with the number of rule firings, as seen in Table 1. Of the 147 rules that existed at the time of the evaluation, 91 fired during the processing of this corpus for a total of 11,929 rule firings. Sixty-nine rules fired five or more times, and 80 were used in more than one message. 62% of all the rules and 57% of the true syntax-semantics mapping rules fired on the test messages. We conclude that, even though the rules were developed by looking at specific examples, they are reasonably general.

Table 2: For the total SCT dataset, the number of syntactic assertions, the number of semantic assertions, and the percent of assertions that are semantic in the syntactic KBs, the semantic KBs, and the semantic KBs without counting the assertions added by CBIR.

	Syntactic	**Semantic**	**Percent Semantic**
Syntactic	2,469	1,149	31.76%
Semantic	538	48,561	98.90%
without CBIR	538	5,646	91.30%

The purpose of the syntax-semantics mapping rules is to convert syntactic information about the words, phrases, clauses and sentences in a message into semantic information about the entities and events discussed in the message, so it is useful to measure the percentage of each KB that consists of semantic assertions. Table 2 shows the number of syntactic assertions[1], the number of semantic assertions, and the percent of assertions that are semantic in the initial syntactic KBs, the final semantic KBs, and the final semantic KBs without counting the semantic assertions added by CBIR (see Section 3.1.3.1.4). The numbers are the totals over all 57 messages of the SCT dataset. As you can see, before the mapping rules, the KBs are almost 70% syntactic, whereas after the mapping rules they are more than 90% semantic. CBIR is purely additive, so it does not reduce the number of syntactic assertions in the KB, but it does increase the semantic content of the KBs to nearly 99%.

[1] The `TokenRange`, `TextOf`, and `RootOf` assertions, which are syntactic, but are retained in the semantic KB for pedigree information and to assist in the downstream scoring of entities against each other, as explained at the end of Section 3.1.3.1.5.3, have been omitted from the

Table 3: Percent of the semantic KBs which are semantic for the BBS and STEF training sets, excluding the CBIR enhancements.

Dataset	Syntactic	Semantic	Percent Semantic
BBS	57	750	92.94%
STEF	517	8,326	94.15%

The percentage of the semantic KBs from the test message set that is semantic, 91.30%, is very similar to that of the training message sets. For example, the semantic content of the semantic KBs of two of these training sets, the BBS and STEF datasets, are 92.94%, and 94.15%, respectively, as shown in Table 3. We conclude that the mapping rules are converting a large part of the syntactic information into semantic information, and doing so in a way that generalizes from the training sets to test sets.

Since the mapping rules were designed using the training datasets, it is possible that some of the rules that fire in our test dataset (as shown in Table 1) are erroneous. That is, the rules may be too general. In order to verify that the rules function as expected, we manually verified that the rules were applied only where they should be.

In order to perform this experiment we ran the mapping rules on each message in the dataset, noting after each rule firing whether the firing was correct or incorrect. Rules which fired due to misparses earlier in the process were not counted as rules used. A rule was counted as firing correctly if its output was semantically valid and in accord with the intent of the rule.

Table 4: The number of rules used in each category, along with the number of times rules from each category were used in the SCT dataset, and the number of times they were used correctly.

Rule Type	Rules Used	Times Fired	Fired Correctly	
			Number	Percent
CBIR	1	474	474	100%
SYN	13	1,567	1,548	98.79%
SEM	5	328	328	100%
SYNSEM	56	2,651	2,431	91.70%
INFER	8	85	72	84.70%
CLEAN	8	6,492	6,492	100%
TOTAL	91	11,597	11,345	97.80%

As Table 4 shows, very rarely were rules applied overzealously. Therefore we can say with some certainty that the rules are not only general enough to fire when processing messages from corpora other than the training set, but they are not overly general; the firings produce a valid semantization of the messages.

Comparison with Other Systems

Our system produces results which are much different from those of the most related system we're aware of—Orbis Technologies' proprietary Cloud Based Text-Analytics (CTA) software. The output of the two systems are not directly comparable. CTA attempts to identify and find relationships among entities, in the process identifying the entities' types as either Person, Organization, Location, Equipment, or Date. Where we identify all the types of entities (and have more types, such as Group and Event), Orbis only seems to identify them when they are in a relation. An Orbis relation is simple— an entity is `associated with` another entity. Tractor uses a large set of relations for representing complex relationships between entities.

Within the 57 SCT messages, Tractor identified (among many other things) 34 entities which were members of specific groups, the religion of 17 entities, 203 locations of events or entities, and 33 persons or groups with specific roles. It additionally identified 102 agents of specific events, 128 themes of events, and over 125 spatial relationships such as "in", "on" and "near".

3.1.3.1.8.2 A Grading Rubric

The evaluation methodology discussed in Section 3.1.3.1.8.1 gives insight into the level of generality of the syntax-semantics mapping rules, and of their thoroughness in converting syntactic information into semantic information. However, it says almost nothing about the correctness of Tractor's semantic analysis of soft information. How is the correctness of a system such as Tractor to be evaluated? For semantic analysis of natural language messages, the notion of "ground truth" does not apply, because regardless of the actual situation being described in the message, if the writer of the message described the situation poorly, no one would be able to reconstruct the situation from the poor description. Instead, the system should be judged by comparing it to a human's performance on the same task. We have developed a scheme for evaluating a message-understanding system by a human "grader" who produces an "answer key," then compares the system's performance to the key [C.1.16].

The answer key is created by the grader's carefully reading the message and listing a series of simple phrases and sentences. The phrases should include all the entities and events mentioned in the message, with the entities categorized into: people; groups of people; organizations; locations; other things, whether concrete or abstract; and groups of things. The simple sentences should express: each attribute of each entity, including the sex of each person for whom it can be determined from the message; each attribute of each event, including where and when it occurred; each relationship between entities; each relationship between events; and each relationship between an event and an entity, especially the role played by each entity in the event. If there are several mentions of some entity or event in the message, it should be listed only once, and each attribute and relationship involving that entity or event should also be listed only once.

If two different people create answer keys for the same message, the way they express the simple phrases and sentences might be different, but even though it might not be possible to write a computer program to compare them, it should still be possible for a person to compare the two answer keys. In this way, a person could grade another person's performance on the message-understanding task. Similarly, if a message-understanding program (e.g., Tractor) were to write a file of entries in which each entry has at least the information contained in the answer key, a person could use an answer key to grade the program.

Tractor writes a file of answers supplying the same kind of entries as the answer key, but with some additional information to help the grader decide when its answers agree with the answer key. For each entity or event other than groups, Tractor lists: a name or simple description; a category the entity or event is an instance of, chosen from the same list given above; a list of the least general categories the entity or event is an instance of; a list of the text ranges and actual text strings of each mention of the entity or event in the message. For each group, Tractor lists: a name or simple description; a category that all members of the group are instances of; a role that all members of the group fill; a list of mentions as above. For each attribute or relationship, Tractor lists an entry in the format `(R a1 a2 ...)`, where `R` is the attribute or relation, a_1 is the entity, group, or event it is an attribute of, or the first argument of the relation, and a_i is the attribute value, or the ith argument of the relation. Excerpts from Tractor's answer file for syn194, in csv format, are:

```
|Dhanun Ahmad|, Person, Person, 16-28:"Dhanun Ahmad" 23-28:"Ahmad" 131-134:"his" 169-172:"his"
(possess |Dhanun Ahmad| detainment58)
(possess |Dhanun Ahmad| Rashid)
(sex |Dhanun Ahmad| male)
Karkh, Location, SectionOfPopulatedPlace, 116-121:"Karkh"
(GeoPosition Karkh wft512: (GeoCoords |33.32174| |44.39384|))
(MGRS Karkh 38SMB4358187120)
cell45, Thing, cell, 108-112:"cell"
(hasLocation cell45 Karkh)
transfer37, Event, transfer, 83-94:"transferred"
(agent transfer37 police32)
(theme transfer37 |Dhanun Ahmad|)
(Date transfer37 |20100303|)
(Time transfer37 time1422)
(recipient transfer37 cell45)
```

Given an answer key, a person can grade another person's answer key, Tractor's submitted answers, or the submission of another message-understanding program. Grading involves comparing the entries in the answer key to the submitted answers and judging when they agree.

We call the entries in the answer key "expected" entries, and the entries in the submission "found" entries. An expected entry might or might not be found. A found entry might or might not be expected. However, a found entry might still be correct even if it wasn't expected. For example, some messages in our corpus explicitly give the MGRS coordinates of some event or location, and MGRS coordinates are also found in the NGA GeoNet Names Server database and added to the KB. If MGRS coordinates were not in the message, but were added, they would not have been expected, but may still have been correct. The grade depends on the following counts: a = the number of expected entries; b = the number of expected entries that were found; c = the number of found entries; d = the number of found entries that were expected or otherwise correct. These counts are combined into evaluation measures adapted from the field of Information Retrieval [C.1.25]: $R = b/a$, the fraction of expected answers that were found; $P = d/c$, the fraction of found entries that were expected or otherwise correct; $F = 2RP/(R + P)$, the harmonic mean of R and P. R, P, and F are all interesting, but F can be used as a summary grade.

3.1.3.1.8.3 Evaluating Tractor's Correctness

We had two undergraduate students, here called "G1" and "G2," create answer keys for the 114 Sunni Criminal Thread (SUN) messages. Each student graded Tractor's performance, and G1 also graded G2's answers. Table 5 shows these grades for the task of identifying the entities and events in the 114 messages. Each average and standard deviation shown is calculated over the 114 messages.

Table 5: Grades of Tractor and a person on identifying entities and events in the 114 SUN messages, showing the average R, P, and F over the 114 messages, and the standard deviations.

Grader	**Performer**	**R**		**P**		**F**	
		Avg	**StD**	**Avg**	**StD**	**Avg**	**StD**
G1	**Tractor**	0.85	0.10	0.87	0.09	0.86	0.08
G2	**Tractor**	0.73	0.12	0.88	0.06	0.79	0.10
G1	**G2**	0.85	0.10	0.71	0.13	0.77	0.11

Notice that grader G1 considered Tractor's performance to be better than G2's (another person). If G2 had graded G1's performance, the assigned R and P should have been the same as G1's P and R, respectively, when grading G2, and the F should have been the same. Thus, G2 would also have considered Tractor's performance to be better than G1's.

Table 6 shows the grades for the task of identifying the entities, events, attributes, and relations in the 114 messages. Again, each average and standard deviation shown is calculated over the 114 messages.

Table 6: Grades of Tractor and a person on identifying entities, events, attributes, and relations in the 114 SUN messages, showing the average R, P, and F over the 114 messages, and the standard deviations.

Grader	Performer	R		P		F	
		Avg	StD	Avg	StD	Avg	StD
G1	Tractor	0.83	0.09	0.83	0.09	0.83	0.08
G2	Tractor	0.66	0.11	0.83	0.08	0.73	0.08
G1	G2	0.79	0.11	0.70	0.14	0.74	0.12

Again, G1 graded Tractor as performing better than G2. G2 would have given Tractor a better P score than G2 would have given G1, and a nearly identical F score, but G2 gave Tractor a particularly low R score.

We conclude that Tractor performs at a human level on the "training" message set.

3.1.3.1.8.4 Evaluating some Components of Tractor [Data being developed]

In order to assess the contributions of some of the components of Tractor, we ran several variants of it, which G1 graded. The variants were:

AutoCoref: Tractor without using the manual co-reference editor;

CycOntology: Tractor with CBIR using Cyc rather than VerbNet/WordNet;

NoGeonet: Tractor without using the GeoNet Names Server;

NoMappingRules: Tractor without the syntax-semantics mapping rules.

G1's grades for these variants of Tractor, as well as for the full Tractor are shown in Table 7.

Table 7: Presents a comparison of the R, P, and F scores of the full Tractor and four variants.

	Entities & Events			Entities, Events, Attributes, & Relations		
	R	P	F	R	P	F
AutoCoref	0.79	0.78	0.78	0.72	0.74	0.73
CycOntology	0.74	0.71	0.72	0.69	0.74	0.71
NoGeonet						

NoMappingRules						
Full Tractor	0.85	0.87	0.86	0.83	0.83	0.83

3.1.3.1.9 CSNePS: A Concurrent Approach to Mapping and Inference

CSNePS is a concurrent implementation of the SNePS 3 knowledge representation and reasoning system [C.1.26] in the Clojure programming language [C.1.28]. SNePS 3 has been used to perform the mapping rules discussed and analyzed above, but it does some work repetitively, and does not take advantage of multi-core computers. The Inference Graphs (IGs) [C.1.27] implemented in CSNePS fix both those problems.

Inference Graphs are built atop propositional graphs – the same graphs used to represent the knowledge in a knowledge base. These graphs are augmented with a prioritized message passing architecture to allow knowledge to flow through the graphs. The priorities are used in concurrent scheduling heuristics to ensure that the knowledge with most usefulness to the current inference operation is given the highest priority. We have found these heuristics to be an order of magnitude better than more naïve approaches. The underlying graph structure allows common parts of rules to be shared, preventing repeated work.

Inference Graphs are more than just a pattern matching mechanism. Unlike the SNePS 3 rule engine, which was built specifically for the mapping rules, compiles rules to machine code, and performs minimal inference, IGs are a full-fledged natural deduction and subsumption inference mechanism. Inference Graphs are capable of performing forward, backward, bi-directional [C.1.29], and focused reasoning [C.2.20] using an expressive first order logic.

CSNePS implements a rule language loosely based on a subset of the syntax of CLIPS [C.1.30], and using concepts from the GLAIR Cognitive Architecture [C.1.31]. Each rule has a name, a left hand side (LHS) and right hand side (RHS). The LHS of a rule is a collection of generic propositions that must be matched (using backward inference) for the rule to fire. The RHS of a rule may contain both Clojure forms and subrules. The set of Clojure forms is executed in order, and the variable bindings from the LHS are substituted in to them appropriately. Rules are implemented as part of an acting system.

In an acting system, a *policy* allows propositions to be connected in some way with *actions*. Actions are often primitive, implemented as code, but using the bindings from the matched propositions. CSNePS rules are implemented as policies, where the LHS contains the propositions to be matched, and the RHS contains the action that should occur. Two rules implemented in the CSNePS rule language are given below.

```
(defrule subjAction
  (nsubj (every action Token Action) (every subj Token))
  =>
  (assert `(~'agent ~action ~subj))
  (unassert `(~'nsubj ~action ~subj)))
```

```
(defrule dobjAction
  (dobj (every action Token Action) (every obj Token))
  =>
  (assert `(~'theme ~action ~obj))
  (unassert `(~'dobj ~action ~obj)))
```

The subjAction rule (seen before in the SNePS 3 rule language in Section 3.1.3.1.6) translates the syntactic relationship of a token which is an instance of Action in an nsubj relationship with another token, into a semantic relation representing that the subject is the agent (performer) of the action. The rule then unasserts the syntactic relationship. The dobjAction rule is very similar to the subjAction rule. It translates the syntactic relationship of a token which is an instance of Action in a dobj (direct object) relationship with another token, into a semantic relation representing that the object is the theme (or, thing undergoing the action). The rule then unasserts the syntactic relationship.

Unlike SNePS 3, which necessarily executes rules one at a time, in CSNePS sets of rules are adopted in a pre-defined order. A set to be adopted may contain just a single rule, or many. When one set completes, the next begins.

One of the major advantages of the graph-based approach used in CSNePS is the ability to share parts of the LHS of rules. The use of shared portions of LHS conditions can be seen by examining the execution time of the two rules described above, subjAction and dobjAction, more carefully (see Table 8). First, in both CSNePS and the SNePS 3 rule engine, we ran just the rule subjAction on all 114 messages of the SUN message set, then we ran both subjAction and dobjAction on those same messages to compare the execution times. The time that the CSNePS IG took to process these two rules was 104% of the time to process subjAction by itself, since much of the LHS of the added dobjAction had already been processed by the system. In SNePS 3, the time to process the two rules was 205% of the time to process the one rule, since the LHS of the rule must be re-processed for every rule, regardless of similarity to other rules already processed. Even though overall CSNePS is slower than SNePS 3 on this test, adding the second rule had less impact in CSNePS both in absolute time, and in percentage of time spent.

Table 8: Time to process the subjAction rule in both CSNePS (using the IG), and the SNePS 3 rule engine, as compared to the time to process both the subjAction and dobjAction rules using those same systems. The difference in time between these two tests shows the advantage of sharing components of the LHS of rules.

Rule Processor	subjAction Time (ms)	subjAction+dobjAction Time (ms)	Time Change
CSNePS IG	78,558	81,413	2,855 (4%)
SNePS 3	4,400	9,000	4,600 (105%)

As mentioned above, SNePS 3 is faster than CSNePS on this task. The same holds true for processing using the mapping rules in general. While this is true, the IG is capable of much more complex inference than SNePS 3, and is designed to be a general tool, able to be used across many domains. The rules we have created on this project were built with the capabilities of the SNePS 3 rule engine in mind – it would be easy to produce rules which CSNePS can handle, but SNePS 3 cannot. Rules have been optimized for the SNePS 3 rule engine, while the same treatment has not been given to the CSNePS versions. We have tested CSNePS on both best-case and worst-case inference tasks, and found that they show a linear speedup with the number of processors. So, it is possible that with enough CPUs dedicated to processing, that CSNePS will out-perform SNePS 3.

3.1.3.2 References

[C.1.1] Michael Prentice, Michael Kandefer, & Stuart C. Shapiro, Tractor: A Framework for Soft Information Fusion, *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*, IFIP, 2010, Th3.2.2, 8 pages, unpaginated.

[C.1.2] H. Cunningham, et al. Text Processing with GATE (Version 6). University of Sheffield Department of Computer Science. 15 April 2011. ISBN 0956599311.

[C.1.3] H. Cunningham, D. Maynard, K. Bontcheva, V. Tablan. "GATE: A framework and graphical development environment for robust NLP tools and applications." Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics (ACL'02). Philadelphia, July 2002.

[C.1.4] Marie-Catherine de Marneffe and Christopher D. Manning. 2008. Stanford Dependencies manual, http://nlp.stanford.edu/software/dependencies_manual.pdf.

[C.1.5] Stuart C. Shapiro, An introduction to SNePS 3. In B. Ganter and G. W. Mineau, Eds., *Conceptual Structures: Logical, Linguistic, and Computational Issues*, Lecture Notes in Artificial Intelligence, vol. 1867, Berlin: Springer-Verlag, 2000, 510-524.

[C.1.6] Daniel R. Schlegel and Stuart C. Shapiro, Visually interacting with a knowledge base using frames, logic, and propositional graphs, In M. Croitoru, S. Rudolph, N. Wilson, J. Howse, and O. Corby, Eds., *Graph Structures for Knowledge Representation and Reasoning*, Lecture Notes in Artificial Intelligence, vol. 7205, Berlin: Springer-Verlag, 2012, 188-207.

[C.1.7] Stuart C. Shapiro and William J. Rapaport, The SNePS family, *Computers & Mathematics with Applications*, vol. 23, no. 2-5, 243-275, January-March 1992. Reprinted in F. Lehmann, Ed., *Semantic Networks in Artificial Intelligence*, Oxford: Pergamon Press, 1992.

[C.1.8] M. Kandefer and S. C. Shapiro, "An F-measure for context-based information retrieval," in Commonsense 2009: Proceedings of the Ninth International Symposium on Logical Formalizations of Commonsense Reasoning, G. Lakemeyer, L. Morgenstern, and M.-A. Williams, Eds. Toronto, CA: The Fields Institute, 2009, pp. 79–84.

[C.1.9] Juan Gómez-Romero, Jesús Garcia, Michael Kandefer, James Llinas, Jose Manuel Molina, Miguel Angel Patricio, Michael Prentice, and Stuart C. Shapiro, "Strategies and Techniques for Use and Exploitation of Contextual Information in High-Level Fusion Architectures", Proceedings of the 13th International Conference on Information Fusion (Fusion2010), Edinburgh, UK. July 2010, TH1.7.3, 8 pages, unpaginated.

[C.1.10] Michael Kandefer and Stuart C. Shapiro, Evaluating Spreading Activation for Soft Information Fusion, *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, 2011.

[C.1.11] Princeton University, "About WordNet," 2010, http://wordnet.princeton.edu.

[C.1.12] M. Palmer, "VerbNet: a class-based verb lexicon," http://verbs.colorado.edu/~mpalmer/projects/verbnet.html.

[C.1.13] B. Levin, English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press, 1993.

[C.1.14] National Geospatial-Intelligence Agency, "NGA GEOnet Names Server," 2014, http://earth-info.nga.mil/gns/html/.

[C.1.15] National Aeronautics and Space Administration, "NASA World Wind," 2011, http://worldwind.arc.nasa.gov/java/.

[C.1.16] Stuart C. Shapiro, A Grading Rubric for Soft Information Understanding, Internal Document, Center for Mutisource Information Fusion, University at Buffalo, December 13, 2013.

[C.1.17] T. Parsons, Events in the Semantics of English: A Study in Subatomic Semantics. Cambridge, MA: MIT Press, 1990.

[C.1.18] Unified Verb Index, University of Colorado, 2012, http://verbs.colorado.edu/verb-index/

[C.1.19] K. J. Miller, "Modifiers in WordNet," in WordNet: An Electronic Lexical Database, C. Fellbaum, Ed. Cambridge, MA: MIT Press, 1998, pp. 47–67.

[C.1.20] S. C. Shapiro, "Symmetric relations, intensional individuals, and variable binding," Proceedings of the IEEE, vol. 74, no. 10, pp. 1354–1363, October 1986.

[C.1.21] K. Sambhoos, J. Llinas, and E. Little, "Graphical methods for real-time fusion and estimation with soft message data," in Proceedings of the 11th International Conference on Information Fusion (Fusion 2008). ISIF, 2008, pp. 1–8.

[C.1.22] J. L. Graham, "A new synthetic dataset for evaluating soft and hard fusion algorithms," in Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011, 2011, pp. 25–29.

[C.1.23] J. L. Graham, J. Rimland, and D. L. Hall, "A COIN-inspired synthetic data set for qualitative evaluation of hard and soft fusion systems," in Proceedings of the 14th International Conference on Information Fusion (Fusion 2011). ISIF, 2011, pp. 1000–1007.

[C.1.24] G. A. Gross, R. Nagi, K. Sambhoos, D. R. Schlegel, S. C. Shapiro, and G. Tauer, "Towards hard+soft data fusion: Processing architecture and implementation for the

joint fusion and analysis of hard and soft intelligence data," in Proceedings of the 15th Int. Conference on Information Fusion (Fusion 2012). ISIF, 2012, pp. 955-962.

[C.1.25] C. van Rijsbergen, Information Retrieval, Second Edition, London: Butterworths, 1979.

[C.1.26] Stuart C. Shapiro, An Introduction to SNePS 3. In Bernhard Ganter & Guy W. Mineau, Eds. *Conceptual Structures: Logical, Linguistic, and Computational Issues*. Lecture Notes in Artificial Intelligence 1867. Springer-Verlag, Berlin, 2000, 510-524.

[C.1.27] Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs. In *Proceedings of the Third International IJCAI Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR 2013)*, 2013, in press.

[C.1.28] Rich Hickey. The Clojure programming language. In *Proceedings of the 2008 Symposium on Dynamic languages*. ACM New York, NY, USA, 2008.

[C.1.29] Stuart C. Shapiro, João P. Martins, and Donald P. McKay. Bi-directional inference. In *Proceedings of the Fourth Annual Conference of the Cognitive Science Society*, pages 90–93, Ann Arbor, MI, 1982. The Program in Cognitive Science of The University of Chicago and The University of Michigan.

[C.1.30] Gary Riley and Brian Dantes. CLIPS reference manual, 2007. http://clipsrules.sourceforge.net/documentation/v630/bpg.pdf.

[C.1.31] Stuart C. Shapiro and Jonathan P. Bona. The GLAIR cognitive architecture. International Journal of Machine Consciousness, 2(2):307–332, 2010. doi: 10.1142/S1793843010000515.

3.1.3.3 Publications

[C.2.1] Juan Gómez-Romero, Jesús Garcia, Michael Kandefer, James Llinas, Jose Manuel Molina, Miguel Angel Patricio, Michael Prentice, and Stuart C. Shapiro, "Strategies and Techniques for Use and Exploitation of Contextual Information in High-Level Fusion Architectures", Proceedings of the 13th International Conference on Information Fusion (Fusion2010), Edinburgh, UK. July 2010, TH1.7.3, 8 pages, unpaginated.

[C.2.2] Michael Prentice, Michael Kandefer, and Stuart C. Shapiro, "Tractor: A Framework for Soft Information Fusion", Proceedings of the 13th International Conference on Information Fusion (Fusion2010), Edinburgh, UK. July 2010, Th3.2.2, 8 pages, unpaginated.

[C.2.3] Stuart C. Shapiro, The Jobs Puzzle: A Challenge for Logical Expressibility and Automated Reasoning. In E. Davis, P. Doherty, and E. Erdem, Eds., *Logical Formalizations of Commonsense Reasoning: Papers from the AAAI Spring Symposium*, Technical Report SS-11-06, AAAI Press, Menlo Park, CA, 2011, 96-102.

[C.2.4] Michael Kandefer and Stuart C. Shapiro, Evaluating Spreading Activation for Soft Information Fusion, *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, 2011.

[C.2.5] Michael Prentice and Stuart C. Shapiro, Using Propositional Graphs for Soft Information Fusion, *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*, 2011.

[C.2.6] Daniel R. Schlegel and Stuart C. Shapiro, Visually Interacting with a Knowledge Base Using Frames, Logic, and Propositional Graphs, *Second IJCAI International Workshop on Graph Structures for Knowledge Representation and Reasoning*, 2011.

[C.2.7] Daniel R. Schlegel and Stuart C. Shapiro, Visually Interacting with a Knowledge Base Using Frames, Logic, and Propositional Graphs. In Madalina Croitoru, Sebastian Rudolph, Nic Wilson, John Howse and Olivier Corby, Eds., *Graph Structures for Knowledge Representation and Reasoning, Lecture Notes in Artificial Intelligence 7205*, Springer-Verlag, Berlin, 2012, 188-207.

[C.2.8] Geoff A. Gross, Rakesh Nagi, Kedar Sambhoos, Daniel R. Schlegel, Stuart C. Shapiro, and Gregory Tauer, Towards Hard+Soft Data Fusion: Processing Architecture and Implementation for the Joint Fusion and Analysis of Hard and Soft Intelligence Data, *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*, IFIP, 2012, 955-962.

[C.2.9] Stuart C. Shapiro, Daniel R. Schlegel, and Michael Prentice, Tractor Manual, Internal Document, Center for Mutisource Information Fusion, University at Buffalo, June 6, 2013.

[C.2.10] Stuart C. Shapiro, A Grading Rubric for Soft Information Understanding, Internal Document, Center for Mutisource Information Fusion, University at Buffalo, December 13, 2013.

[C.2.11] Stuart C. Shapiro, and Daniel R. Schlegel, Natural Language Understanding for Soft Information Fusion, *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013),* IFIP, July, 2013, 9 pages unpaginated.

[C.2.12] Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs (student abstract), *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, AAAI Press/The MIT Press, Menlo Park, CA, 2013, 1637-1638.

[C.2.13] Daniel R. Schlegel, Concurrent Inference Graphs (Doctoral Consortium abstract), *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence (AAAI-13)*, AAAI Press/The MIT Press, Menlo Park, CA, 2013, 1680-1681.

[C.2.14] Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs. In *Proceedings of the Third International IJCAI Workshop on Graph Structures for Knowledge Representation and Reasoning (GKR@IJCAI 2013)*, 2013, unpaginated, 9 pages.

[C.2.15] Daniel R. Schlegel and Stuart C. Shapiro, Inference Graphs: A Roadmap. In Matthew Klenk and John Laird, Eds. *Proceedings of the Second Annual Conference on Advances in Cognitive Systems*, 2013 Poster Collection, December, 2013, 217-234.

[C.2.16] Daniel R. Schlegel and Stuart C. Shapiro, Concurrent Reasoning with Inference Graphs. In Madalina Croitoru, Sebastian Rudolph, Stefan Woltran, and Christophe Gonzales, Eds., *Graph Structures for Knowledge Representation and Reasoning, Lecture Notes in*

Artificial Intelligence 8323, Springer International Publishing, Switzerland, 2014, 138-164. DOI: 10.1007/978-3-319-04534-4_10

[C.2.17] Geoff A. Gross, Ketan Date, Daniel R. Schlegel, Jason J. Corso, James Llinas, Rakesh Nagi, and Stuart C. Shapiro, Systemic Test and Evaluation of a Hard+Soft Information Fusion Framework: Challenges and Current Approaches, *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*, IFIP, July, 2014, unpaginated, 8 pages.

[C.2.18] Daniel R. Schlegel and Stuart C. Shapiro, Inference Graphs: A New Kind of Hybrid Reasoning System (student abstract), *Proceedings of the Twenty-Eighth Conference on Artificial Intelligence (AAAI-14)*, AAAI Press/The MIT Press, Menlo Park, CA, 2014, in press.

[C.2.19] Schlegel, D.R. and Shapiro, S.C., Inference Graphs: A New Kind of Hybrid Reasoning System, *Proceedings of the Cognitive Computing for Augmented Human Intelligence Workshop at AAAI-14 (CCAHI@AAAI-14)*, 2014, 4 pages, in press.

[C.2.20] Daniel R. Schlegel and Stuart C. Shapiro, The `Ah Ha!' Moment : When Possible, Answering the Currently Unanswerable using Focused Reasoning. In P. Bello, M. Guarini, M. McShane, & B. Scassellati, Eds., *Proceedings of the 36th Annual Conference of the Cognitive Science Society (COGSCI 2014)*, Cognitive Science Society, Austin, TX, 2014, 1371-1376.

[C.2.21] Stuart C. Shapiro and Daniel R. Schlegel, Natural Language Understanding for Information Fusion. In Galina Rogova and Peter Scott, Eds., *Fusion Methodologies in Crisis Management Higher Level Fusion and Decision Making for Crisis Management*, Springer S. A., in press

[C.2.22] Michael Kandefer and Stuart C. Shapiro, Context Relevance for Text Analysis and Enhancement --- Soft Information Fusion. In Lauro Snidaro, Jesus Garcia, Erik Blasch and James Llinas, Eds., *Boosting Real World Performance with Domain Knowledge,* Springer S. A., forthcoming.

[C.2.23] Daniel R. Schlegel, Concurrent Inference Graphs, PhD Dissertation, Department of Computer Science and Engineering, University at Buffalo, Buffalo, New York, 2014, forthcoming.

3.1.4 Graph Analytic Processing

3.1.4.1 Stochastic Graph Matching

The graph analytic techniques (batch/incremental stochastic graph matching [D.1]) developed in years 1 and 2 went through extensive testing in year 3 to verify both their ability to produce optimal graph matching results and compare the runtime to existing techniques. To evaluate these performance metrics batch graph matching executions of the stochastic truncated search tree approach (TruST) were compared with a math programming formulation solved with the commercial mixed integer solver CPLEX version 12.4 [D.2]. The mathematical formulation is as follows:

$$x_{ij} = \begin{cases} 1 \text{ if data graph node } v_i \text{ is matched with tempalte graph node } v_j \\ 0 \text{ otherwise} \end{cases}$$

$$y_{iujv} = \begin{cases} 1 \text{ if data graph edge } e_{iu} \text{ is matched with template graph edge } e_{jv} \\ 0 \text{ otherwise} \end{cases}$$

$S(v_i, v_j)$ = similarity score between data graph node v_i and template graph node v_j

$S(e_{iu}, e_{jv})$ = similarity score between data graph edge e_{iu} and template graph edge e_{jv}

Objective function:

$$\max_{x_{ij}, y_{iujv}} \left\{ \sum_{v_i} \sum_{v_j} S(v_i, v_j) x_{ij} + \sum_{e_{iu}} \sum_{e_{jv}} S(e_{iu}, e_{jv}) y_{iujv} \right\}$$

Subject to:

$$\begin{aligned}
&x_{ij} \geq y_{iujv} && \forall e_{iu} \in E_D, e_{jv} \in E_T \\
&x_{uv} \geq y_{iujv} && \forall e_{iu} \in E_D, e_{jv} \in E_T \\
&\textstyle\sum_{v_i} x_{ij} = 1 && \forall v_j \in V_T \\
&\textstyle\sum_{v_j} x_{ij} \leq 1 && \forall v_i \in V_D \\
&\textstyle\sum_{e_{iu}} y_{iujv} = 1 && \forall e_{jv} \in E_T \\
&x_{ij} \in \{0, 1\} && \forall v_i \in V_D, v_j \in V_T \\
&0 \leq y_{iujv} \leq 1 && \forall e_{iu} \in E_D, e_{jv} \in E_T
\end{aligned}$$

Due to the fuzzily defined similarity scores, their corresponding ranking values were utilized in the objective function. Randomly generated template graphs ranging in size from 3 to 6 nodes were used in the testing. A pilot study was run on independent data to determine appropriate TruST run parameters. Using the pilot study generated parameters the top 10 CPLEX generated solutions were compared with the top 10 TruST solutions both based on solution runtime and quality. The results are shown in .

Table 9, Table 10, Figure 12 and Figure 13.

Table 9: TruST Speedup over CPLEX at 10 Results

TruST Speedup over CPLEX (10 Results)

		Data Graph Node Count					
		500	1000	2000	3000	4000	5000
TG Node Count	3	9.53	14.84	10.87	12.23	16.03	20.56
	4	12.68	16.17	38.52	35.49	46.41	51.55
	5	4.14	5.42	15.22	22.98	22.98	29.21
	6	5.70	13.46	26.29	48.00	59.43	60.77

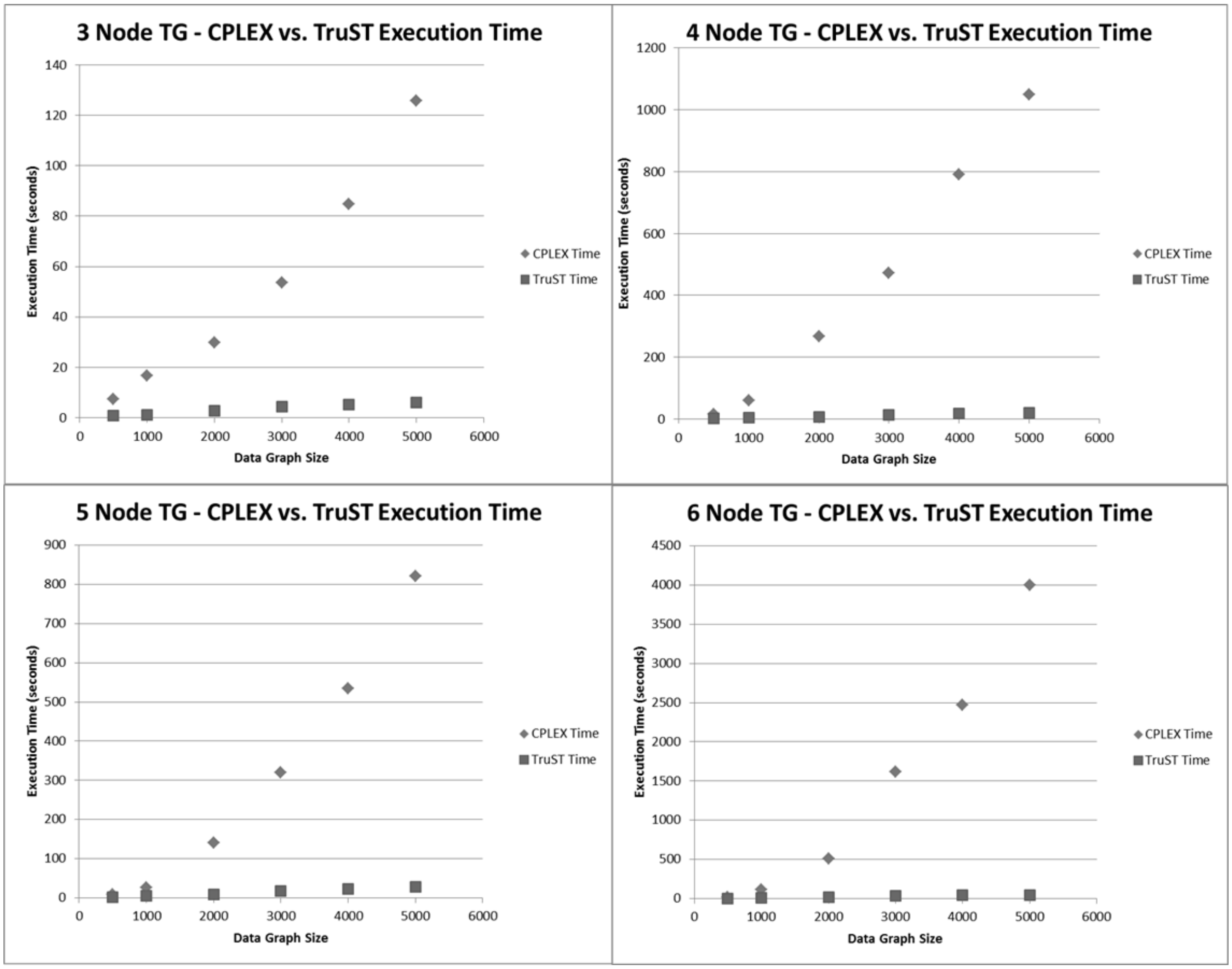


Figure 12: CPLEX vs. TruST Execution Time Graphs

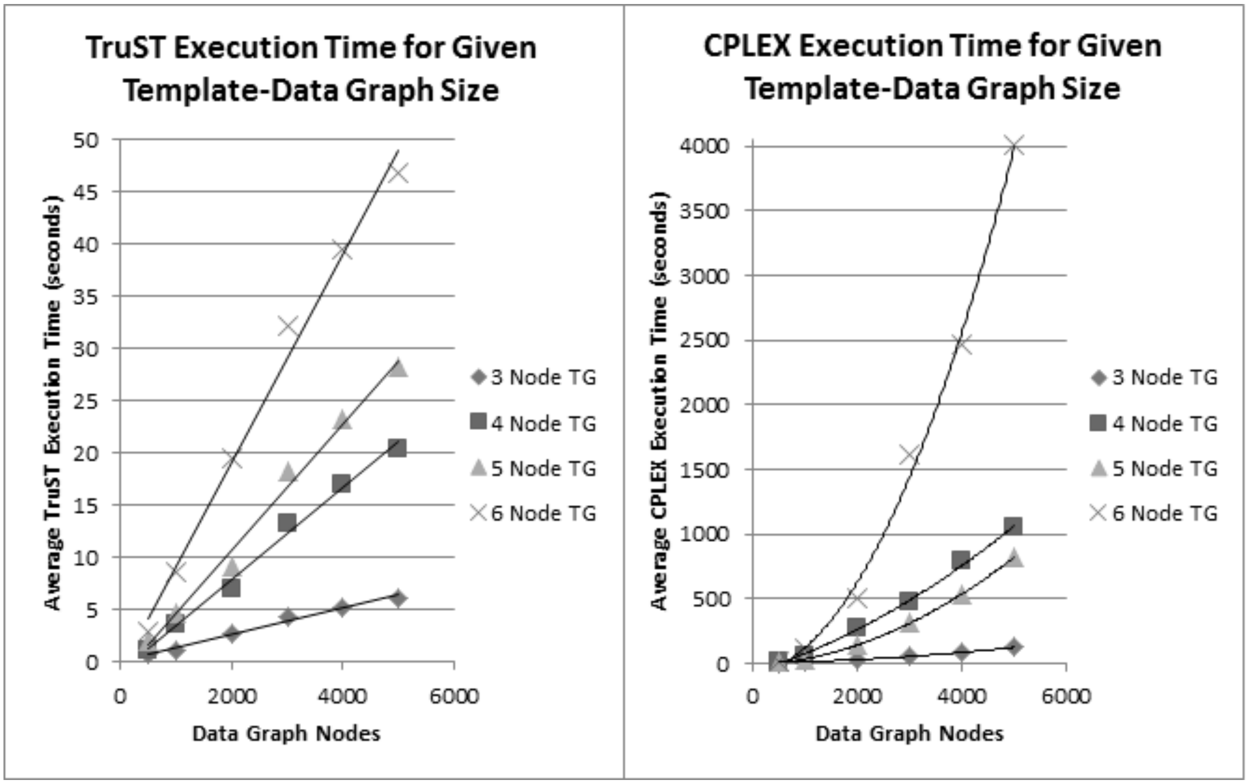


Figure 13: Execution Time versus TG-DG Size

Table 10: Average Optimality Gap (top 10 solutions)

Average Optimality Gap (10 Results)

		Data Graph Node Count					
		500	1000	2000	3000	4000	5000
TG Node Count	3	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	4	0.04%	0.00%	0.01%	0.00%	0.18%	0.34%
	5	0.00%	0.00%	0.00%	0.00%	0.00%	0.11%
	6	0.00%	0.00%	0.00%	0.00%	0.23%	0.07%

From the results we see that the average CPLEX execution time grows roughly quadratically in the size of the data graph, while the average TruST execution time grows linearly. In addition to the scalability benefit it is also shown that the overall time averaged speedup of the TruST algorithm is 42.7 (i.e., CPLEX execution takes on average 42.7 times more time to identify 10 results than TruST execution). These results are obtained while maintaining a minimal (≤ 0.34%) optimality gap under all experimental conditions, with 17 of 24 experimental conditions having no optimality gap.

3.1.4.1.1 Incremental Matching Numerical Studies

The incremental graph matching algorithm was tested versus the batch algorithm with a specific interest in quantifying the speedup provided by the incremental algorithm. The speedup is evaluated both including the time required to reconstruct the existing search tree results and ignoring this time. Including the time required to reconstruct the search tree emulates an environment where the search tree cannot be maintained in memory (and thus must be rebuilt at each incremental invocation). If there were available memory to persist the search tree, the execution times without search tree reconstruction are more representative of that use case. The test replicates were run on initial graph sizes of 2,000 nodes each with 2,000 increments added. Combined results from the test replicates are shown in Table 11, Figure 14 and Figure 15.

Table 11: Speedup of incremental algorithm over batch matching execution for Various experimental settings

	Template Graph Node Count			
	3	4	5	6
Incremental Speedup with Search Tree Reconstruction	36.2	44.5	33.9	36.9
Incremental Speedup without Search Tree Reconstruction	46.4	60.8	41.5	42.9

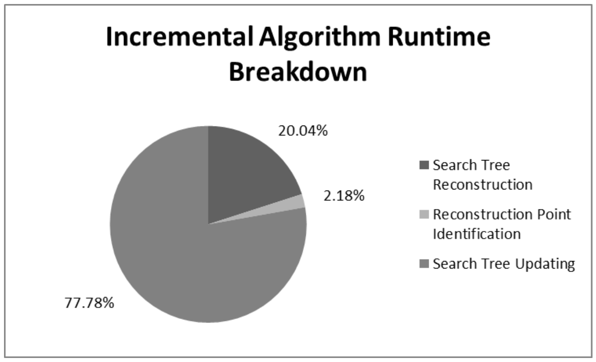


Figure 14: Incremental Algorithm Runtime Breakdown by Activity

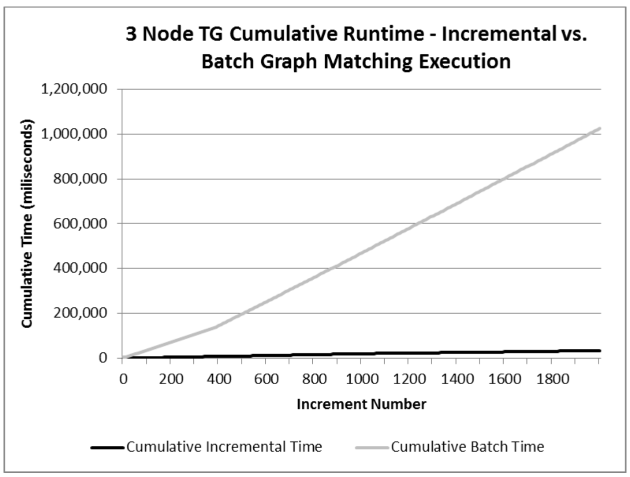


Figure 15: Cumulative Runtime Example by Increment Number (Averaged over 10 Replicates)

3.1.4.1.2 AND/OR Stochastic Graph Matching

The intelligence analysis processes considered here seek to facilitate the interaction between a military commander and all-source intelligence analyst. In the intelligence process the commander provides information requirements (IR) which the intelligence analysts must answer. The IR are defined as: "Those items of information regarding the enemy and his environment which need to be collected and processed in order to meet the intelligence requirements of a commander" [D.3]. These IR are prioritized by the commander to indicate priority information requirements (PIR) which are defined as IR which will influence the overall success of a mission. PIR are prioritized among each other and the ordering can change over the course of an operation [D.3],[D.4]. These ranked IR serve as the input to the intelligence analyst's workflow. Given the ranked set of IR, the analysts break them down into specific information requirements (SIR) which make up some portion of an IR. Each IR is decomposed into multiple SIR [D.4]. These SIR are themselves made up of a series of indictors and warnings which signal the existence of the higher level IR.

When considering an application of the graph matching algorithm in an applied environment, considerations for the simultaneous identification of multiple IR must be made. An analyst is not simply searching for and maintaining an awareness of the existence of a single IR and its related

template graph(s). The purpose of a template graph is to establish an awareness of the degree to which a complex situation exists in the observed domain. In the domain of intelligence analysis these matches are meant to answer Priority Intelligence Requirements (PIR). PIR can be responded to most effectively when broken down into a collection of SIR and indicators or warnings. The utilization of a report based synthetic dataset for counter-insurgency (SYNCOIN, [D.5]) combined with PIR and indicators provided by a potential commander, lead to the identification that many PIR indicators and thus their template graph representations contain common elements. This realization paired with the potential for increased algorithmic efficiency and the presence of topological errors transmitted from upstream processing elements form the motivation for AND/OR template graphs.

An example indicator which an intelligence analyst would attempt to identify is as follows: "Threats against public works, utilities or transportation. Threats of violence against prominent personalities." Notice the indicator is composed of threats against any one of the listed potential targets. When represented as a series of template graphs, the indicator section representing the threat is required regardless of a target of public works, utilities, etc. A template graph representation of this indicator is seen in Figure 16. PIR provide higher level assessments than indicators, combining multiple indicators with similar AND and OR relationships. Thus, to assess PIR while maintaining an efficient solution method, PIR templates are constructed recursively with a "node" in a PIR template representing a single indicator (e.g., public works).

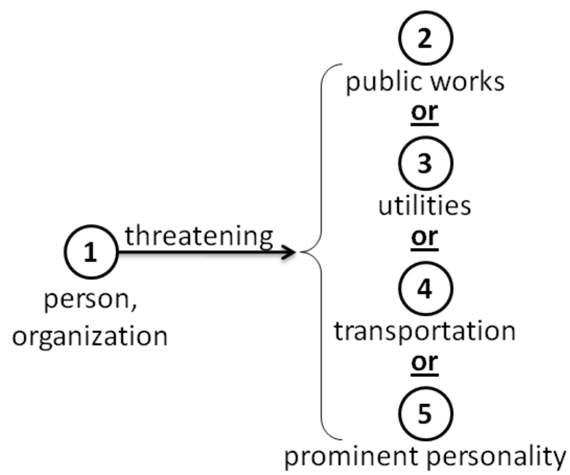


Figure 16: Example indicator template

The template graphs previously suggested for use in stochastic graph matching (developed up through Year 3 in [D.6]) required the existence of each template graph node to form a valid template graph to evidential graph match. In an AND/OR template graph this requirement is relaxed to require only the template graph nodes on a single AND path be present in the template graph to evidential graph match. A simple example will be used to illustrate this concept.

An example AND/OR template graph is displayed in Figure 16. In this graph Node 1 is required while either Node 2, 3, 4 or 5 complete the graph. The four AND paths in this graph are: Nodes 1 and 2, Nodes 1 and 3, Nodes 1 and 4, and Nodes 1 and 5. To avoid multiple graph matching executions with redundant branchings on Node 1 the AND/OR structure of the templates is utilized. The methodology for matching AND/OR template graphs is described subsequently.

The extensions of the existing stochastic graph matching algorithm ([D.6]) to allow for an AND/OR structured template graph require some method for conveying the AND and OR relationships. We will begin with some terminology employed throughout this section. An "AND path" is a series of graph elements which are connected by AND logical relations, meaning each of the elements is required for a match to that path. In the template graph displayed in Figure 16, Nodes 1 and 2 and the connecting edge form a single AND path. A graph matching result which matches all nodes and edges on a particular AND path constitutes a complete template graph to evidential graph match. These paths may be of different lengths depending on the form of the graph.

An "OR set" is a set of graph elements in which a match of any one of the set members satisfies the requirements of the set. In the template graph displayed in Figure 16, Nodes 2, 3, 4 and 5 make up an OR set. It should be noted that members of an OR set are not limited to a particular size, do not have to contain the same number of graph elements and can be recursively defined.

The topological structure of an AND/OR template graph is the same as that of an AND template graph. However, in addition to the topology of the graph, the OR set relationships within the graph must be made clear to the matching algorithm. These OR set relationships are represented through the use of a precedence tree.

3.1.4.1.2.1 Precedence Tree

The precedence tree specifies the allowable branching order during graph matching execution. In the case of an AND/OR template graph the precedence tree is built with the goal of branching on the most common graph elements first (i.e., graph elements existing in the most AND paths) with the hope of minimizing the redundant branchings which would be expected of individual templates for each AND path.

Additional motivation for the use of a precedence tree structure in the matching of AND/OR templates is provided by the inclusion of precedence tree branch specific beam width proportions. These precedence tree branch specific beta allocation parameters control the proportion of the beam width which is reserved for that solution path. These parameters can be set with different objectives in mind. Some objectives which might be considered include template subsection salience consideration, maximizing solution variety, maximizing solution quality (irrespective of AND path) or weighting AND Paths based on relative importance. The details of the precedence tree design principals and beta allocation settings are omitted here. The interested reader can see Sections 3.2.1 and 3.2.2 of [D.7] for these details.

3.1.4.1.2.2 Numerical Testing

Evaluation of the previously described graph matching methods involves comparison of both solution quality and runtime. In the following section we perform three main comparisons: 1.) comparing the performance of a single AND path template graph with and without a template precedence tree specified, 2.) comparison of individual templates without precedence information to a single AND/OR template and 3.) the comparison of individual templates with precedence information to a single AND/OR template. The random data for each experimental treatment was generated as realistically as possible. The topology was randomly generated from actual social network data with entity types and attributes generated from census data.

As previously mentioned we are interested in both the solution runtime and quality. While runtime quantification is straightforward, the definition of result "quality" is more difficult. The analyst is interested in the investigation of varied results with high similarity scores, not simply the best match. This interest stems from the possibility of multiple graph regions fulfilling IR with differing levels of similarity/completeness. To help capture this interest in the evaluation of the matching algorithms it is insufficient to compare only the single highest similarity graph matching solution. A similarity score percentage deviation quality evaluation metric is used in the place of the traditional optimization solution quality metric of optimality gap. This metric compares the solution quality among alternate graph matching runs for a predefined number of top solutions.

The solution quality evaluation metric can be evaluated in two different ways depending on the objective of the AND/OR matching run. If the objective of the AND/OR matching run was to identify top solutions for each of the AND Paths this metric can be evaluated across the top n solutions for each AND Path (i.e., $n * |AND\ Path|$ solutions are considered where $|AND\ Path|$ is the number of AND Paths for the AND/OR template graph). However, if the AND Paths within the template graph represent the same situation we may only be interested in the top n solutions irrespective of which AND Path they were found on. In this case the evaluation metric will be calculated on these top n solutions, ignoring the AND Path which they came from.

In addition to a solution quality metric the evaluation metric of algorithmic speedup is considered when comparing graph matching results of an AND/OR template graph versus multiple AND Path template graphs. The speedup is calculated as the cumulative runtime for each AND Path template graph divided by the runtime of the corresponding AND/OR template graph. A speedup above 1 indicates a runtime benefit of the AND/OR template graph over the individual AND Path template graphs. Finally, a concept of "AND/OR Efficiency" is introduced to compare the relative effectiveness of different AND/OR graph topologies. AND/OR efficiency is calculated as the AND/OR speedup divided by the number of AND Paths contained within that AND/OR template graph.

The comparison between individual templates and an AND/OR template is performed in two ways. Individual templates are matched both with and without precedence specification. These results are then compared to the single AND/OR template, evaluating the metrics described earlier in this section. The experimental factor levels (and in fact the data and template graph content) remain the same between the individual templates with and without precedence specification, with each experimental treatment run for each of the precedence tree beta objectives. For each experimental setting 10 replicates were run in which unique data graph and template graph attribute representations were generated. Data graph sizes considered include 500, 1000 and 2000 nodes while the AND/OR template graph topologies utilized are pictured in Figure 17 (Note: Edge direction may be switched to maintain the appropriate directed relationships based on the randomly generated entity types). There are a number of search tree run parameters which must be set depending on the experimental setting. A pilot study was run to identify the appropriate algorithm parameters.

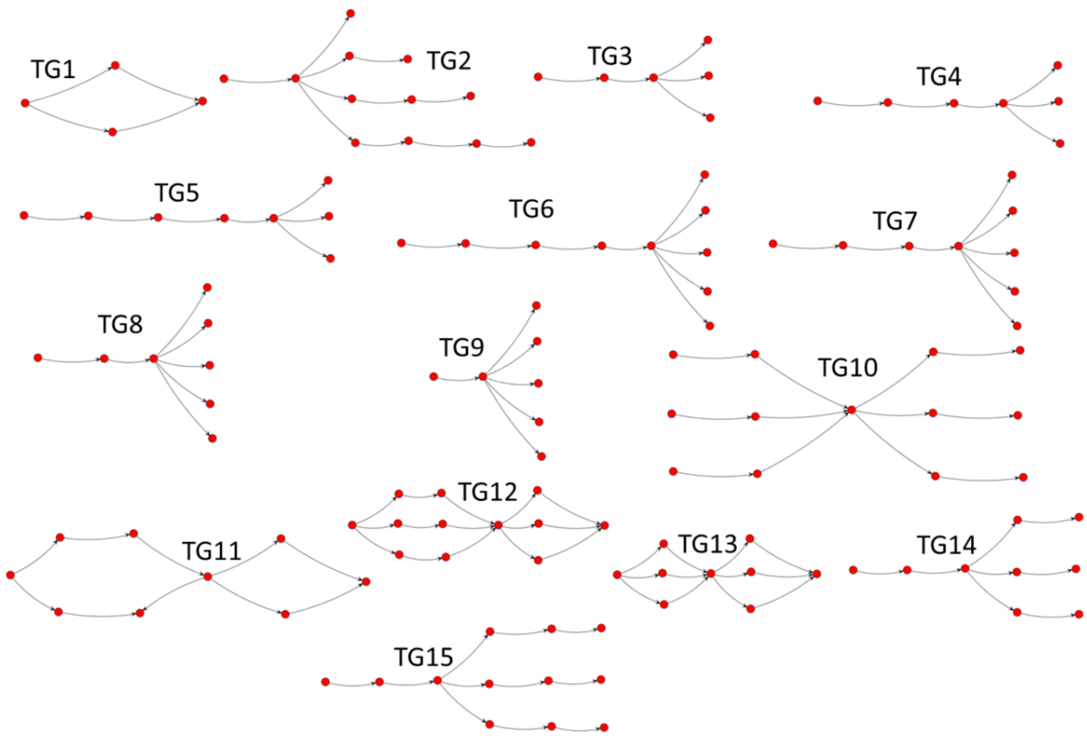


Figure 17: Template graph (TG) topologies utilized in numerical testing

3.1.4.1.2.3 AND/OR Graph Matching - Numerical Testing

Here we consider the beta objective of maximizing solution variety. In this objective we are interested in finding top solutions for each of the AND Paths specified by an AND/OR template graph. The solution quality evaluation metric is evaluated across the top 5 results for each AND Path. Three types of templates are considered in this comparison: a single AND/OR template graph, individual AND path template graphs (1 per AND path of the corresponding AND/OR template graph) and individual AND path template graphs which include the same precedence tree as the AND/OR template graph (Note only the results between a single AND/OR template graph and individual template graphs without precedence specification are presented here, other results can be seen in [D.7]).

The results comparing an AND/OR template graph to individual templates for each AND Path are shown in

Table 12, Table 13, and Table 14. From these results we see a significant runtime improvement through the use of a single AND/OR template graph at the expense of some minor loss in solution quality. The AND/OR efficiency is improved as the degree of overlap between AND Paths increases. For example, the AND/OR efficiency increases from 0.66 to 0.90 to 1.17 between template graphs 3, 4 and 5 respectively. Notice the only change to the template graph topology is the addition of another AND required node from template graph 3 to 4 and 4 to 5.

Table 12: Average speedup of AND/OR template graphs over individual templates

	TG Number															
DG Size (Nodes)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	DG Size Avg.
500	1.36	1.66	2.10	2.89	3.56	4.44	3.29	2.46	1.52	3.66	3.77	6.09	5.17	2.71	3.13	3.19
1000	1.40	1.36	1.93	2.66	3.55	4.24	3.23	2.13	1.63	3.38	3.12	5.59	4.72	2.41	2.89	2.95
2000	1.26	1.32	1.91	2.59	3.44	4.60	2.79	1.84	1.42	3.42	3.24	5.50	5.26	2.35	2.42	2.89
TG Number Avg.	**1.34**	**1.44**	**1.98**	**2.71**	**3.51**	**4.43**	**3.11**	**2.14**	**1.52**	**3.49**	**3.38**	**5.73**	**5.05**	**2.49**	**2.81**	**3.01**

Table 13: Solution quality gain of AND/OR template graphs over individual templates

	TG Number															
DG Size (Nodes)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	DG Size Avg.
500	-1.18%	0.00%	-2.15%	-0.07%	-0.33%	-0.17%	-0.01%	-0.09%	-0.21%	-1.10%	-3.39%	-3.55%	-1.67%	-0.33%	0.38%	-1.15%
1000	-1.87%	-0.02%	0.00%	-0.20%	-0.17%	-0.65%	-0.27%	0.00%	-0.12%	-2.05%	-0.32%	-0.80%	-0.10%	-0.19%	-0.86%	-0.57%
2000	-0.45%	-0.01%	-0.20%	-0.30%	-0.14%	-0.08%	0.00%	0.00%	-0.15%	-0.65%	-0.08%	-1.11%	-0.22%	-0.09%	-0.53%	-0.33%
TG Number Avg.	**-1.17%**	**-0.01%**	**-0.78%**	**-0.19%**	**-0.21%**	**-0.30%**	**-0.09%**	**-0.03%**	**-0.16%**	**-1.27%**	**-1.27%**	**-1.82%**	**-0.66%**	**-0.20%**	**-0.34%**	**-0.68%**

Table 14: AND/OR Efficiency by template graph (averaged over DG sizes) versus individual templates

TG Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AND Path Count	2	4	3	3	3	5	5	5	5	9	4	9	9	3	3
Average Speedup	1.34	1.44	1.98	2.71	3.51	4.43	3.11	2.14	1.52	3.49	3.38	5.73	5.05	2.49	2.81
AND/OR Efficiency	0.67	0.36	0.66	0.90	1.17	0.89	0.62	0.43	0.30	0.39	0.84	0.64	0.56	0.83	0.94

An AND/OR efficiency of over 1 is possible due to the more restrictive nature of the AND/OR template graph. The additional restriction of branching on AND required nodes initially leads to fewer search tree branchings (the count of which provide a good predictor of matching runtime). For example, one replicate of template graph 5 for a data graph size of 2,000 nodes required only 6,948 search tree branchings in the case of the AND/OR template while 60,060 search tree branchings were performed for the corresponding AND path template graphs. This significant difference in the number of branchings did not result in any solution quality gap at 5 solutions per AND Path.

Other results compare AND/OR stochastic graph matching to individual template graphs with identical precedence tree specification, a different methodology of one-hop neighborhood score calculation, expansion of the AND/OR search tree's beam width, similarity score profiling based beta allocation and dynamic (online) beta allocation. These results are omitted here but can be seen in [D.7].

Past graph matching approaches fail to take advantage of overlapping situations of interest in a multi-template environment, requiring multiple graph matching executions to identify these potentially overlapping situations. The methodology presented here of precedence tree guided search and AND/OR graph matching enables the ability of matching multiple situations of interest with a single graph matching execution. This single graph matching execution is shown to provide significant speedup over multiple single template matching executions, in many cases approaching a speedup near the number of simultaneous situations of interest being matched.

This increased algorithmic efficiency is shown to come at little solution quality loss. Search tree breath allocation techniques are shown to obtain desired solution quality/variety tradeoffs while identifying matches to the AND/OR template graphs. Additional extensions of AND/OR neighborhood scoring, score profiling aided search tree breadth allocation and dynamic search tree breadth allocation are also shown to be beneficial under certain conditions.

3.1.4.1.3 Link Analysis

The identification of paths between key nodes of a graph is an important problem in a number of domains as well as a subproblem of many other graph analytic techniques. While the identification of paths may seem trivial, practical difficulties with latencies in graph traversal and work in progress data size explosion must be overcome when working with very large graphs. The problem considered here is the execution of a link analysis query.

Link analysis identifies in a graphical data store the connections between two or more entities of interest (EOI). These EOI may be of a variety of types (e.g., persons, locations, organizations, etc.) and may be directly connected or connected via a long chain of relationships (e.g., a direct communication or distant familial relationship respectively). In an unconstrained computational environment (unlimited memory) a breath first search (BFS) approach can be utilized in the identification of paths. Due to the scale of the data store and potential for exponential explosion of intermediary nodes in the number of hops away from an EOI, the sequential BFS approach is not feasible in this environment. The approach implemented here (partially a Year 3 effort) is a scalable, parallel breath first search within the Apache MapReduce framework.

The link analysis implementation is analogous to a parallel breadth first search, moving one hop (edge) away from the current node at each iteration of the algorithm. Some terminology utilized throughout the remainder of this section is as follows:

- **Entity of Interest (EOI)** – any entity for which we are interested in identifying paths to and from other EOI
- **Root Node** – a particular instance of an EOI within the considered graph; the starting point for the algorithmic branching (NOTE: There may be more than one root node corresponding to a single EOI if de-duplication/data association has not been performed)
- **Sub-path** – a partial path consisting of nodes (at least one of which is an EOI) and edges; expanded upon at each algorithm iteration in attempt to locate another sub-path coming from a different EOI

The following are the main algorithmic steps:

0. Identify instances of the EOI within the global graph
1. Branch from the existing frontier (root nodes only for first iteration)
2. Determine overlap of newly reached nodes with paths from other root nodes
   a. Merge and output connected paths if any were formed at this iteration
      i. Output single merged path for input at next iteration
   b. Output non-connected paths as input for next iteration
3. Check termination criteria, if not met, return to Step 1
   a. Possible Termination Criteria

i. Number of iterations
ii. Number of solutions
iii. Runtime
iv. Analyst interrupt

Step 0 identifies one or more instance of each EOI within the background knowledgebase, providing the root nodes for the iterative branching. Step 1 expands the current paths (only root nodes at first iteration) by one hop via both outgoing and incoming edges. Step 2 determines if any of the expanded paths (from Step 1) have either met at a common node or crossed parallel edges. If paths have met, they are merged and output for analyst and next iteration consideration (discarding sub-paths). Paths, which have not met, are output as the input for the next iteration. Step 3 checks if any of the termination criteria have been met, returning to the branching step if the algorithm should continue.

3.1.4.1.3.1 Link Analysis Algorithm Example

The following example demonstrates the execution of the link analysis algorithm within the Hadoop framework, providing an illustration of each <Key, Value> pair throughout the programs execution.

The following example assumes the analyst is interested in the connections between nodes 1, 2 and 3. The sub-paths beginning at each EOI are represented by yellow, blue and green nodes respectively (see Figure 18).

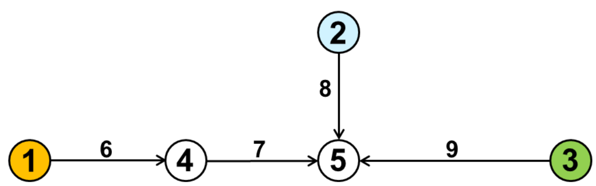


Figure 18: Step 0: EOI identification

<Key, Value> pairs representing these root nodes form the input to the first iteration of the Hadoop job. The first mapper moves one hop from these root nodes, outputting the one-hop sub-paths (see Figure 19).

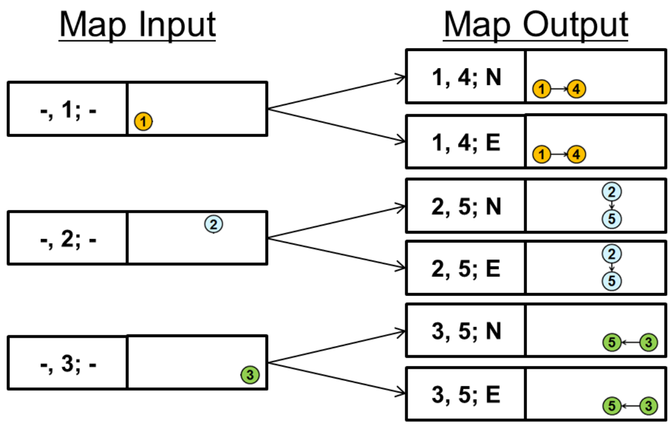


Figure 19: Iteration 1: Step 1. branch from root nodes

The map output is processed by a partitioner which determines which of the available reducers the <Key, Value> pair should go to. The partitioner considers the key alone when determining which reducer the value should go to. In the case of a node searching key the partitioner determines the partition number based on the last node (ensuring paths which are currently at the same node end up at the same reducer). If the key is edge searching, the partitioner ensures values which have the same last two nodes (in any order) end up at the same reducer. As seen in Figure 20, node searching values 3 and 5 arrive at the same reducer call (Reducer 3) where the paths are merged (see the connected path consisting of EOI 2 and 3 and intermediary node 5). After reduction, the merged values are output for node searching keys only. This output also serves as the input for the next link analysis iteration if no termination criterion is met. By only outputting values from node searching keys we ensure branching effort is not duplicated at the next iteration.

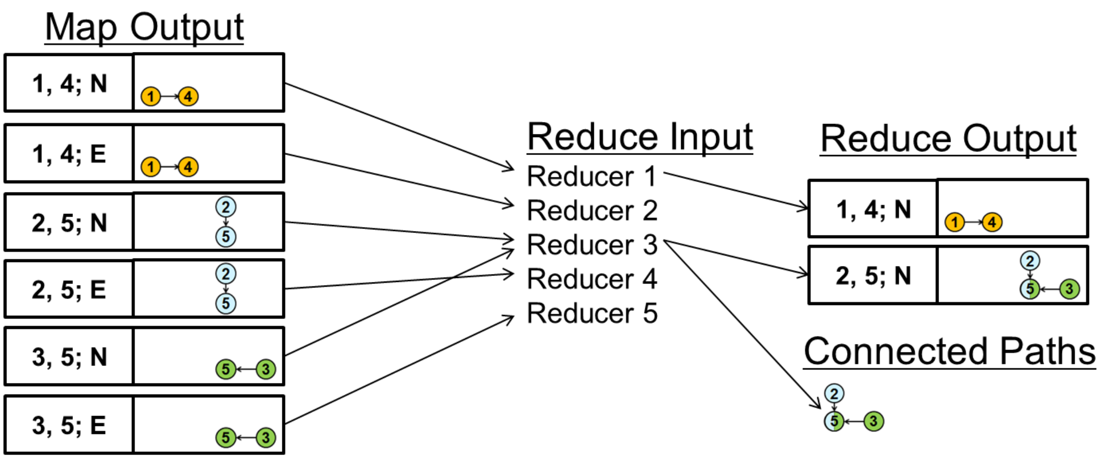


Figure 20: Iteration 1: Step 2. Partitioning and reduction (merging)

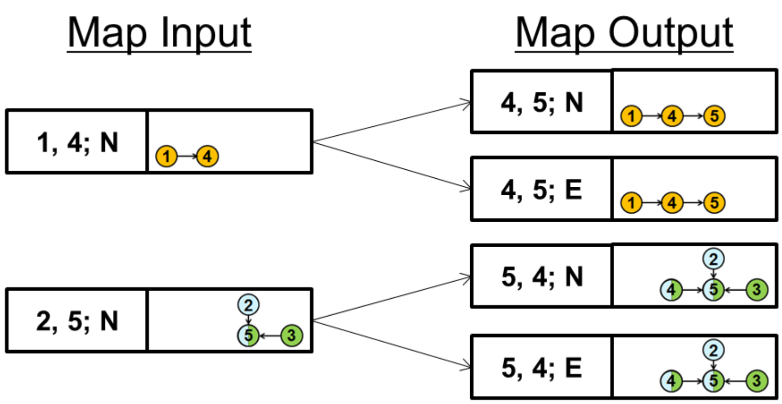


Figure 21: Iteration 2: Step 1. Branch from iteration 1 reduce output

Note that the second partitioning step moves edge searching values (<Key, Value> pairs 2 and 4) to the same reducer (Reducer 2). These edge searching keys have visited the same last two nodes (Nodes 4 and 5), meaning they have crossed over a common edge or set of edges if there is more than one edge connecting the nodes. These sub-paths are merged at the reducer and the connected path output for analyst consideration.

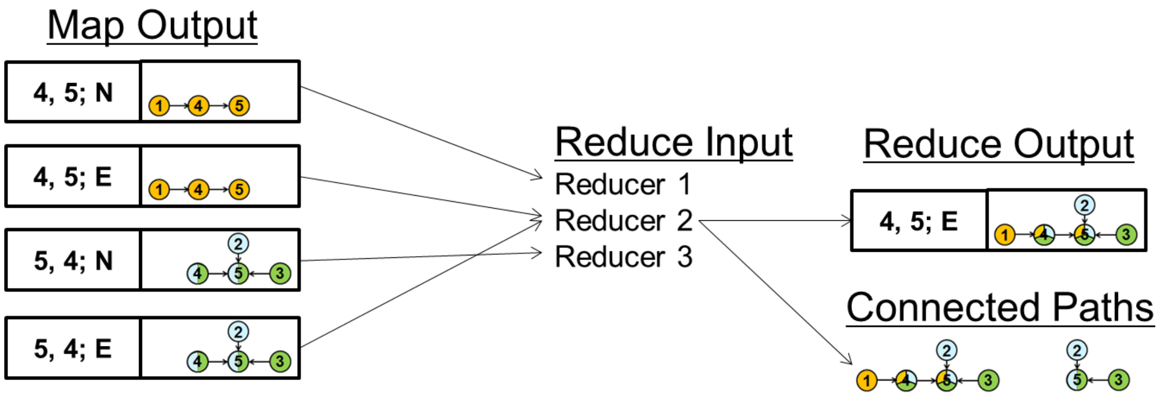


Figure 22: Iteration 2: Step 2. Partitioning and reduction (merging)

Potential termination criteria which would end the link analysis program execution after the second iteration include: a two iteration limit, a connected path limit of 2, a runtime limit which was passed in the second iteration or an analyst interrupt.

3.1.4.1.3.2 Link Analysis - Numerical Testing

Numerical testing was performed at the University at Buffalo Center for Computational Research (CCR) [D.8]. Hadoop clusters of varying size were dynamically allocated for testing of algorithm scalability. The nodes utilized in these experiments were Dell E5645 2.4 Ghz 12 core nodes each with 48 GB of memory. These nodes are networked via an Ethernet connection. Each node also serves as an HBase data node in the tests of the HBase data access methodology. A dedicated PostgreSQL server is configured on a separate 12 core node which has 48 GB of RAM dedicated to the PostgreSQL instance.

The test data set consists of 10 million entities and 19 million edges, each with attributes, totaling 20 GB in raw TSV form. Random link analysis queries are drawn from the 10 million entities with 100, 200 and 300 starting points making up the test set. Five instances exist of each starting point count for a total of 15 test link analysis queries.

A summary of the number of adjacencies requested for each query by iteration is shown in

Table 15. The explosion of intermediate data is exemplified by these results as indicated by large number of adjacencies requested at the 3rd iteration.

Table 15: Adjacency requests by query and iteration number

Adjacency Requests by Query and Iteration Number

		Iteration Number		
		1	2	3
Query Number	1	100	230	22,129
	2	200	471	43,531
	3	300	777	66,001
	4	100	232	18,044
	5	200	523	47,601
	6	300	700	53,892
	7	100	251	19,174
	8	200	509	36,489
	9	300	753	64,046
	10	100	263	19,484
	11	200	515	49,456
	12	300	757	63,608
	13	100	270	22,619
	14	200	517	41,349
	15	300	792	67,573

A number of cluster configurations were tested with an interest in identifying the capacity of the data access methodologies to support the algorithm data requirements. Cluster sizes of 2 and 8 nodes were considered with between 2 and 12 tasks (mappers and reducers) simultaneously running per node. Average algorithmic runtime by data access method, iteration, cluster node count and task count is shown in Figure 23. The results of these trials show the improvement of the Postgres data access method until the number of simultaneous connections becomes too large, resulting in database thrashing (around the 64 task point). Meanwhile HBase improves with diminishing returns with the addition of concurrent tasks. The ideal settings for the Postgres connection occur with 8 nodes and 32 tasks with an average algorithm runtime of 85.3 seconds over the 15 trial link analysis jobs. The ideal HBase settings are 8 nodes and 64 tasks with an average algorithm runtime of 97.1 seconds (see Figure 23).

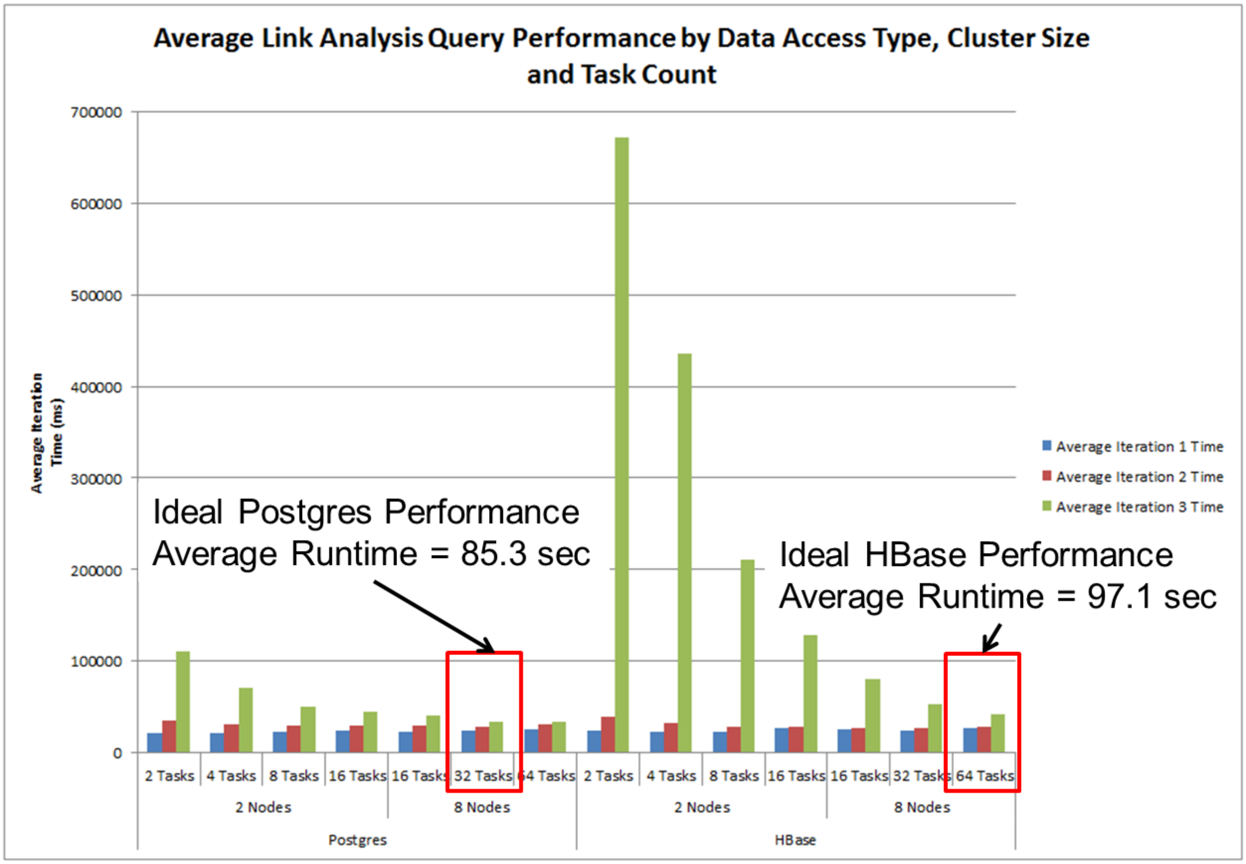


Figure 23: Ideal performance cluster configurations

The driving force in the algorithmic runtime is the speed of adjacency requests. Table 16 identifies the average adjacency retrieval time by data access method and iteration under ideal cluster configuration settings for that data access method. As indicated by the average link analysis job execution time, Postgres data access is significantly faster than HBase in the most time consuming iteration, iteration 3.

Table 16: Average adjacency retrieval time by data access method and iteration under optimal cluster settings

Adjacency Retrieval Time by Data Access Method and Iteration

		Iteration 1	Iteration 2	Iteration 3
Data Access Method	HBase	30.58	31.00	11.09
	Postgres	7.38	44.17	2.43

The average and average maximum mapper and reducer task times by data access method, iteration number and task count are shown in Table 17 and Table 18 respectively. The maximum times are more indicative of the algorithm runtime since the reducer start time (ignoring data transfer to the reducer) is blocked by waiting for the final mapper to finish. Also, the next iteration mapper is blocked by waiting for the previous iterations last reducer to finish. From these average maximum times we see that under the ideal (overall) settings of 8 nodes and 32 and 64 tasks for Postgres and HBase data access respectively, Postgres is slightly faster in

iteration 1, HBase is faster in iteration 2 and Postgres is significantly faster in iteration 3. The significant maximum mapper time in iteration 3 is the reason Postgres outperforms HBase overall.

Another point to note from Table 18 is the domination of the algorithm runtime by the mapping phase. Under the ideal cluster configuration the mapper consumes 86% of the runtime for Postgres data access and 90% of the time for HBase data access.

Table 17: Average mapper/reducer task times (in ms) by data access method, iteration number and task count

Average Mapper/Reducer Task Times by Data Access Method, Iteration Number and Task Count

		Average Mapper Times											
		Iteration 1				Iteration 2				Iteration 3			
	Task Count	16	32	64	92	16	32	64	92	16	32	64	92
Data Access Method	HBase	1991	1579	2040	2112	2166	1840	2153	2401	32331	17848	10449	9017
	Postgres	1382	1127	1794	2231	2615	2141	2264	2323	7696	4802	3890	3968

		Average Reducer Times											
		Iteration 1				Iteration 2				Iteration 3			
	Task Count	16	32	64	92	16	32	64	92	16	32	64	92
Data Access Method	HBase	481	498	584	741	794	699	697	781	1140	939	853	911
	Postgres	469	423	493	511	886	682	668	762	1201	1049	963	861

Table 18: Average maximum mapper/reducer task times (in ms) by data access method, iteration number and task count

Average Maximum Mapper/Reducer Task Times by Data Access Method, Iteration Number and Task Count

		Average Maximum Mapper Times											
		Iteration 1				Iteration 2				Iteration 3			
	Task Count	16	32	64	92	16	32	64	92	16	32	64	92
Data Access Method	HBase	2756	2383	3150	3304	3942	3318	3501	4259	66123	37249	21738	19212
	Postgres	1837	1703	2587	3221	7726	5176	5961	5866	18123	11100	9978	9695

		Average Maximum Reducer Times											
		Iteration 1				Iteration 2				Iteration 3			
	Task Count	16	32	64	92	16	32	64	92	16	32	64	92
Data Access Method	HBase	674	728	989	1454	972	953	1033	1355	1440	1173	1274	1587
	Postgres	615	614	869	940	1145	891	1039	1444	1530	1358	1591	1434

Additional tests were performed to assess the impact of data replication in HBase. These tests replicated the data twice for each of the 8 node task settings. While this did considerably reduce network congestion it did not prove to improve link analysis runtime.

3.1.4.2 References

[D.1] Gross, Geoff A., Rakesh Nagi, and Kedar Sambhoos. "A Fuzzy Graph Matching Approach in Intelligence Analysis and Maintenance of Continuous Situational Awareness." *Information Fusion,* (in revision August 2012)

[D.2] CPLEX 12.4, http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/

[D.3] United States. Department of the Army. FM 101-5-1 Operational Terms and Graphics. 1997.

[D.4] United States. Department of the Army. FM 34-130 Intelligence Preparation of the Battlefield. 1994.

[D.5] Graham, Jacob L., David L. Hall and Jeffrey Rimland, "A synthetic dataset for evaluating soft and hard fusion algorithms." Proc. SPIE 8062, 80620F (2011).

[D.6] Gross, Geoff A., Rakesh Nagi, and Kedar Sambhoos, "A Fuzzy Graph Matching Approach in Intelligence Analysis and Maintenance of Continuous Situational Awareness." Information Fusion, (accepted May 2013).

[D.7] Gross, Geoff A., Rakesh Nagi, "Precedence Tree Guided Search for the Efficient Identification of Multiple Situations of Interest – AND/OR Graph Matching." Journal of Heuristics, (submit June 2013).

[D.8] Center for Computational Research. <http://ccr.buffalo.edu/>.

3.1.5 Systemic Testing of a Hard+Soft Information Fusion System

A topic which has recently received much attention within the information fusion domain is the topic of Hard+Soft information fusion. Hard+Soft information fusion considers both hard, physical sensor (e.g., radar, acoustic, etc.) and soft, linguistic (e.g., human reports, Twitter feeds, etc.) data sources. Many modern domains both in the military and private industry settings (e.g., counterinsurgency [E.1],[E.2], disaster relief [E.3],[E.4], consumer marketing [E.5],[E.6], etc.) have come to recognize the importance of the fusion of numerous data sources, broadly including both hard and soft data. One research effort which has confronted the subject of hard+soft fusion is the Multi-disciplinary University Research Initiative (MURI) on Network-based Hard+Soft Information Fusion [E.7].

The MURI program in Hard+Soft Information Fusion has developed a fully integrated hard+soft fusion research prototype system in which raw hard and soft data are processed through hard sensor processing algorithms (e.g., detection and tracking), natural language understanding processes, common referencing, alignment, association and situation assessment fusion processes. The MURI program is currently in its 5th year. During years 1 through 4, the MURI team dealt with research issues in developing a baseline hard+soft fusion system, while identifying a number of design alternatives for each of the framework processing elements. A recent focus (to continue through program completion) is in the systemic test and evaluation (T&E) of the developed hard+soft information fusion framework.

While traditional experimental or training approaches may be used in assessing processes of a hard+soft information fusion framework in isolation, the nature of dependencies across framework components requires a systemic approach in which the cross-component affects are understood. Past efforts in the T&E of hard, soft and hard+soft information fusion systems have

largely focused on the evaluation of situational awareness of the human or machine consumer of system output (e.g., [E.8], [E.9], [E.10], [E.11]). While this assessment is an important measure of system effectiveness,[2] these past studies generally do not include assessments of sub-process performance and its effect on overall system performance (i.e., producing an error audit trail). In this paper we describe the design of a metric-centric test and evaluation framework for systemic error trail analysis and parametric optimization of hard+soft fusion framework sub-processes. We will discuss the performance metrics utilized including notions of "system optimality," issues in defining the parametric space (design variants), cross-process error tracking methodologies and discuss some initial results.

The remainder of this paper is structured as follows: Section 3.1.5.1 defines the exemplar system under test and Section 3.1.5.2 describes issues in defining metrics and the parametric space (or system variants) to be considered within T&E. Section's 3.1.5.3-3.1.5.7 provide an overview of the framework processes within the exemplar system under test and provide both individual process and cross-process evaluation metrics. Specifically, Section 3.1.5.3 introduces one physical sensor processing element within the MURI framework (as an exemplar of T&E approaches for these hard data processes), Section 3.1.5.4 presents an overview of the natural language understanding evaluation methodology, Section 3.1.5.5 describes the system benefit of the common referencing process (uncertainty alignment), Section 3.1.5.6 explains the evaluation of the data association process (readers are directed to [E.12] for a more detailed description) and Section 3.1.5.7 identifies a variety of graph analytic techniques which are applied on the cumulative associated data to enable situation assessments. Finally, Section 3.1.5.8 discusses some initial T&E results across these framework processes and plans for future work and Section 3.1.5.9 provides conclusions.

3.1.5.1 System Under Test (SUT)

A necessity when performing system T&E is the definition of a System Under Test (SUT), which is the set of functional components and connections to be evaluated. The definition of the SUT must consider the larger project schedule beyond the T&E efforts. For example, continuing research and development (R&D) work during the T&E period may make the SUT a moving target. A decision may need to be made whether to freeze the SUT or allow for the continuing evolution of framework processes (see Section 3.1.5.2 for additional thoughts on tracking SUT performance through R&D iterations). Particularly if R&D efforts are to continue throughout the T&E period, version control and version logging must be carefully followed such that results and process settings of any test run may be replicated.

While the methods and metrics developed in this paper are fairly general, we will consider specific applications to the system architecture developed within the MURI project [E.1] (see Figure 24). Within the MURI framework, hard (or physical sensor) input data enters the hard sensor fusion and track creation processes which convert the raw sensor data (video, acoustic, etc.) into semantic tracks, containing the entity and attribute evolution over the duration of the data and some interaction events. Evaluation of the hard sensor fusion processes is described in Section 3.1.5.3.

[2] Although "situational awareness" provides a measure of the degree to which the system supports user understanding, many systems require further support, and an assessment of the degree to which the system facilitates action on this obtained understanding. Not much work toward this higher level objective exists within the literature and this topic is noted for a direction of future work.

Soft (or linguistic) input data within the MURI framework enters the Tractor Natural Language Understanding (NLU) process which performs processes including: dependency parsing, within-source co-reference resolution, named entity identification, morphological analysis to find token root form, context-based information retrieval and syntax-semantics mapping. The resulting propositional graph from Tractor is ideally fully semantic content (versus syntactic), containing all of the semantic propositions which would be identified by a human interpreter of the original message. Evaluation of this capability is described in Section 3.1.5.4.

After a conversion from a propositional graph to attributed graph, the soft data stream is run through a common referencing and uncertainty alignment process. This process seeks to account for observational biases and variances in human observation, accounted for by contextually-based human error models, developed within this program. Evaluation of the benefit of this process to the fusion tasks of data association and situation assessment is described in Section 3.1.5.5.

Next, the hard and soft data streams enter the data association process. Data association algorithmically identifies common entities, events and relationships across data sources and data modalities, associating the entities, attributes, and relationships based on computed similarity criteria. The objective of data association is to form a single node for each unique entity or event or a single edge for each unique relationship within the cumulative data (see Section 3.1.5.6).

Upon the formation of a cumulative, fused body of evidence (the cumulative associated data graph), analyst-guided graph analytic processes reason over this data in an attempt to obtain and maintain situational estimates. Some graph analytic processes which were developed under the MURI effort (along with initial evaluation considerations) are described in Section 3.1.5.7.

3.1.5.2 Defining Metrics and the Test Space

With the SUT defined, a determination of evaluation points within the SUT must be made. The evaluation points within the MURI SUT are separable along process lines including: physical sensor processing, natural language understanding, data association and graph analytic processes (situation assessment) as shown in Figure 24. For each of these processes we define evaluation metrics which are expected to be reflective of overarching system performance. Potential performance metrics are broadly classified as quality and runtime-based metrics, with the simultaneous optimization of both typically resulting in a conflicting objective. Depending on the operational environment, solution quality or runtime may be at a premium. Due to the basic research nature of our program and lack of a specific target data environment, our focus was on quality-based metrics.

While the physical sensor and natural language understanding processes operate on raw data which is expected to be factually correct,[3] downstream processes of data association and graph analytics may be subject to upstream errors. As a result, these downstream processes must consider the notion of both process and cumulative system optimality. The performance metrics for each process are described in detail within Section's 3.1.5.3-3.1.5.7.

[3] We assume the hard and soft data streams contain factual information not resulting from intentional attempts to deceive. While we understand these data (in particular soft data) may be subject to contradictions, inconsistencies or deception, the resolution of these elements was not a focus or expectation of the MURI program.

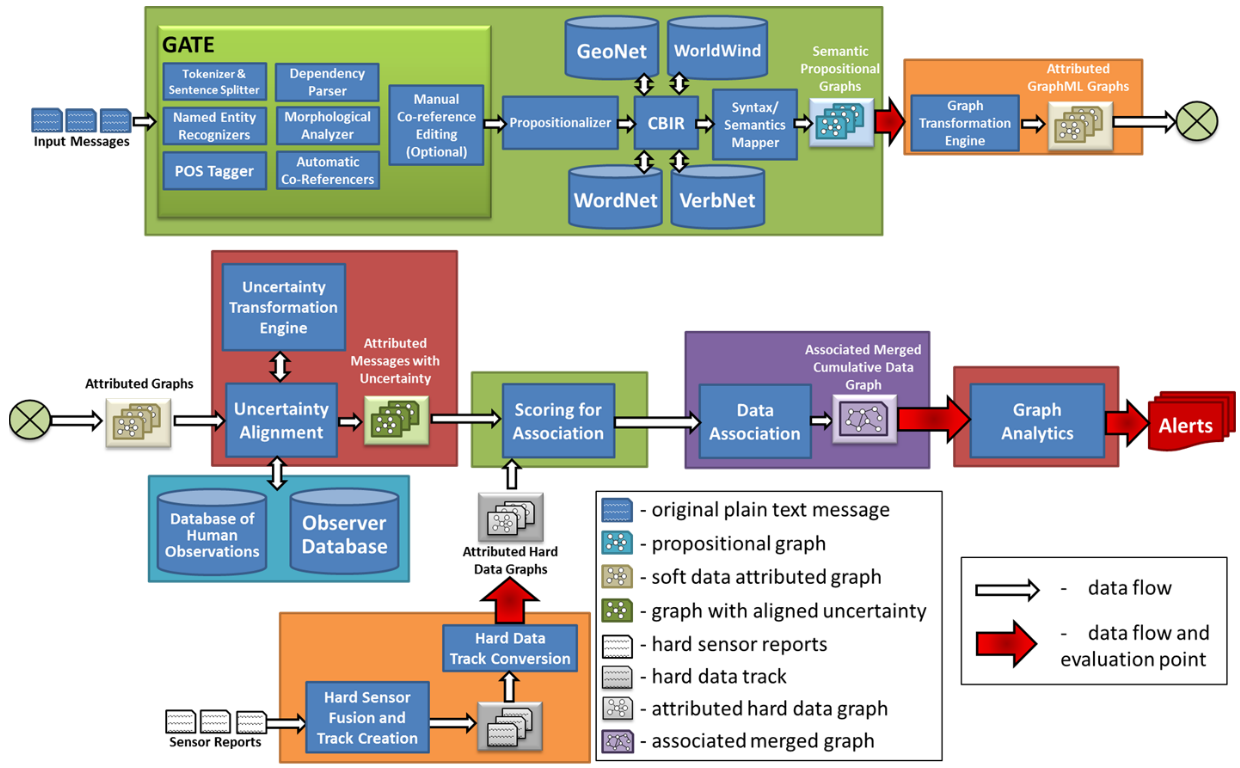


Figure 24: MURI System Under Test

In addition to identifying a system configuration resulting in "system performance optimality," another overarching goal of system T&E is to measure the main effects and interactions of design alternatives on both process and system-level performance metrics. While the number of design alternatives which could be considered is theoretically infinite (e.g., numerical parameters), some pilot study or process expert guidance may be used to prune the potential training and evaluation space. In addition to utilizing identified performance metrics as a basis for spiral (incremental) system development, a number of experimentation questions were developed, thus defining a *test space* for experimentation.

In addition to process parameters, elements of the test space include input data qualities. A natural interest in the nascent area of hard+soft information fusion is the quantification of the value of hard versus soft versus hard+soft information to some system level objective (e.g., to situational awareness performance measures). An additional input data interest within the test space is the robustness of processes to varied levels of input data quality, whether raw data or machine processed. The assessment of situational awareness metrics after the graph analytic processes in our SUT remains as future work (see Section 3.1.5.7).

In addition to the optimization of each of the many process parameters, a sampling of process variation questions to be assessed via the T&E processes described subsequently are as follows:

1. How general are each of the processes to variations in input data? What are the input data qualities which affect system performance?
2. What is the effect of alternate stemmers within the NLU process?

3. How do different ontologies used within NLU processing (and downstream processes) affect performance?
4. How robust is the data association process to variations in input data quantity and quality?
5. What is the ideal recall/precision tradeoff in data association to best support situational awareness at the graph analytic processes?

The metrics identified in support of the evaluation of the above experimental questions are described subsequently.

3.1.5.3 Physical Sensor Tracking and Attribution Evaluation

We use a Deformable Part Model [E.13], abbreviated DPM, to detect specific instances of object categories in the hard data video frames. The DPM method is the state-of-the-art object detection method in the computer vision literature [E.14]; it depends heavily on methods for discriminative training and combines a margin-sensitive approach for data mining hard negative examples within a formalism called latent SVM (Support Vector Machine). The DPM model represents an object as a set of parts that are permitted to locally displace (translate; despite the name deformable, there is no actual deformation in the model) allowing it to adapt to variations in object structure, articulations, and weak visual evidence. The model uses histograms of oriented gradients [E.15] as local features extracted from the images. During inference, the parts are allowed to displace locally and the reported detection score is the one that yields a maximum score over all configurations of the local parts.

To facilitate fair experimentation on the relatively small SYNCOIN physical sensor dataset (see Figure 25), we directly used the car and the human (upright pedestrian) DPM models that are available in the software package from Felzenswalb's PASCAL VOC experiments (see [E.13]). In other words, we do not train a separate DPM model specifically in our experimental scenario because the available samples are too few. The Felzenswalb's PASCAL VOC models are trained on the respective PASCAL VOC data, which are images and not video. Performance improvements are expected if trained on domain-specific data.

For tracking after detection, we use a tracking-by-detection framework and dynamic programming to compute best-fit tracks over the videos [E.16]. The basic method computes a best-fit path through the full set of detected objects over time. The best-fit minimizes a deformation penalty (penalizes large frame-to-frame motion) and computes the globally optimal tracks for the given set of detected objects.

For evaluation of the hard data extraction, we rely on well-established techniques from the computer vision community PASCAL VOC benchmark [E.14]. Specifically, for each semantic category, such as vehicles and people, we conduct a separate evaluation. Since we are concerned with detection, we essentially evaluate "for a given image, where are the instances of category X (if any)?" As in the PASCAL VOC, we will use the average precision metric to evaluate the detections. The first part of the evaluation is determining a positive hit for which we use the intersection-over-union criterion. Following PASCAL VOC, let the predicted bounding box for a given task be denoted by B_p and the ground truth be denoted by B_g. We compute an overlap ratio: $\rho = \frac{area(B_p \cap B_g)}{area(B_p \cup B_g)}$. When the overlap threshold exceeds a predetermined value (PASCAL VOC suggest 0.5) then the detection is considered a positive hit.

Given these positive hits, for average precision of a given task and class, we compute the standard precision-recall curve. The average precision is used to compute a summary statistic of the shape of the precision-recall curve. It is computed as the mean precision for a uniformly spaced set of recall values. The PASCAL VOC uses eleven such recall values, and we will follow this specification.

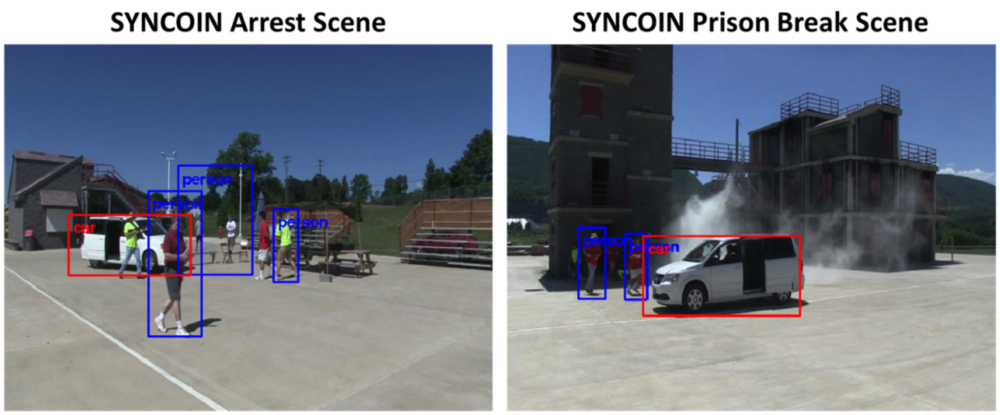


Figure 25: Example detections on the SYNCOIN videos.

3.1.5.4 Natural Language Understanding Evaluation

Tractor [E.17],[E.18] is the subsystem of our hard+soft fusion system that is designed to understand soft information. In this context, understanding soft information means creating a knowledge base (KB), expressed in a formal knowledge representation (KR) language, that captures the information in an English message. Tractor operates on each message independently, and outputs a formal KB consisting of a series of assertions about the situation described in the message. The assertions include the categories (or types) each entity and event mentioned in the message is an instance of, the attributes of those entities and events, and the relations among the entities and events. The assertions are expressed in the SNePS 3 KR language [E.19],[E.20] and can be viewed as forming a propositional graph [E.21]. The assertions that are extracted from the message are enhanced with relevant ontological information from VerbNet [E.22] and WordNet [E.23] and geographical information from the NGA GeoNet Names Server database [E.24].

How is a system such as Tractor to be evaluated? Within Tractor, the notion of "ground truth" does not apply, because regardless of the actual situation being described in the message, if the writer of the message described the situation poorly, no one would be able to reconstruct the situation from the poor description. Instead, the system should be judged by comparing it to a human's performance on the same task. We present a scheme for evaluating a message-understanding system by a human "grader" who produces an "answer key," then compares the system's performance to the key.

The answer key is created by the graders carefully reading the message and listing a series of simple phrases and sentences. The phrases should include all the entities and events mentioned in the message, with the entities categorized into: people; groups of people; organizations; locations; other things, whether concrete or abstract; and groups of things. The simple sentences should express: each attribute of each entity, including the sex of each person for whom it can be determined from the message; each attribute of each event, including where and when it occurred; each relationship between entities; each relationship between events; and each relationship between an event and an entity, especially the role played by each entity in the event. If there are

several mentions of some entity or event in the message, it should be listed only once, and each attribute and relationship involving that entity or event should also be listed only once.

If two different people create answer keys for the same message, the way they express the simple phrases and sentences might be different, but even though it might not be possible to write a computer program to compare them, it should still be possible for a person to compare the two answer keys. In this way, a person could grade another person's performance on the message-understanding task. Similarly, if a message-understanding program (e.g., Tractor) were to write a file of entries in which each entry has at least the information contained in the answer key, a person could use an answer key to grade the program.

Tractor writes a file of answers supplying the same kind of entries as the answer key, but with some additional information to help the grader decide when its answers agree with the answer key. For each entity or event other than groups, Tractor lists: a name or simple description; a category the entity or event is an instance of, chosen from the same list given above; a list of the least general categories the entity or event is an instance of; a list of the text ranges and actual text strings of each mention of the entity or event in the message. For each group, Tractor lists: a name or simple description; a category that all members of the group are instances of; a role that all members of the group fill; a list of mentions as above. For each attribute or relationship, Tractor lists an entry in the format (R a_1 a_2 ...), where R is the attribute or relation, a_1 is the entity, group, or event it is an attribute of, or the first argument of the relation, and a_i is the attribute value, or the i^{th} argument of the relation.

Given an answer key, a person can grade another person's answer key, Tractor's submitted answers, or the submission of another message-understanding program. Grading involves comparing the entries in the answer key to the submitted answers and judging when they agree. We call the entries in the answer key "expected" entries, and the entries in the submission "found" entries. An expected entry might or might not be found. A found entry might or might not be expected. However, a found entry might still be correct even if it wasn't expected. For example, some messages in our corpus explicitly give the MGRS coordinates of some event or location, and MGRS coordinates are also found in the NGA GeoNet Names Server database and added to the KB. If MGRS coordinates were not in the message, but were added, they would not have been expected, but may still have been correct. The grade depends on the following counts: a = the number of expected entries; b = the number of expected entries that were found; c = the number of found entries; d = the number of found entries that were expected or otherwise correct. These counts are combined into evaluation measures adapted from the field of Information Retrieval [E.25]: $R = b/a$, the fraction of expected answers that were found; P$= d/c$, the fraction of found entries that were expected or otherwise correct; $F = 2RP/(R + P)$, the harmonic mean of R and P. R, P, and F are all interesting, but F can be used as a summary grade. Average grades for 80 messages of the SYNCOIN dataset are, R=0.83, P=0.84, F=0.83.

3.1.5.5 Common Referencing and Uncertainty Alignment

We consider the common referencing process of uncertainty alignment [E.26],[E.27]. Uncertainty alignment attempts to resolve a number of inconsistencies within the soft data stream including: qualitative language (e.g., "tall" person), human observational biases and variance and uncertainty transformations if required (e.g., enabling comparisons between fuzzy and probabilistic uncertainty representations). Due to the uncertain nature of inferences made by the uncertainty alignment process, it is difficult to quantify these results as "correct" or

"incorrect." As a result, within our T&E of the uncertainty alignment process (see [E.26]) we have assessed the benefit of uncertainty alignment to the fusion processes of data association and situation assessment (through graph matching). This T&E process has shown a significant benefit of uncertainty alignment to both data association and graph matching.

3.1.5.6 Data Association

3.1.5.6.1 Overview

If hard+soft data sources contain duplicate references to the same real world entity, event or relationship, the data association process needs to be performed, for merging common entities, events and relationships into fused evidence. This fused evidence is used in sense-making processes to make inferences on the state of the real world (obtain situational awareness) [E.1]. The data association problem can be modeled as a graph association problem. Different data association formulations (Graph Association or GA^N, Multidimensional Assignment problem with Decomposable Costs or MDADC and Clique Partitioning Problem or CPP) and their related algorithms for data association were studied on this program by Tauer et al. [E.28],[E.29] and Tauer and Nagi [E.30], each of which has its own strengths and weaknesses.

The first step of data association is to measure and quantify the similarity between pairs of nodes (or edges) in the input dataset. These similarity scores are calculated using a similarity function, which provides a positive score if two elements are similar; and a negative score if two elements are dissimilar. The absolute value of the similarity score is an indication of the strength of similarity or dissimilarity between a certain node/edge pair.

Given these similarity scores, data association tries to cluster (or associate) the nodes/edges which are highly similar, and produces a *cumulative data graph* (CDG), which is the cumulative fused evidence. The cumulative evidence should describe the real world as accurately as possible from the provided input data, so as to draw satisfactory conclusions on the state of the real world. This calls for the development of an objective strategy for training and evaluating the performance of data association processes. This evaluation strategy also needs to be efficient with minimal human intervention. In this section, we will briefly describe the evaluation methodology that has been developed for assessing data association both with a "system perspective" and isolated "data association perspective."

3.1.5.6.2 Evaluation Methodology

The evaluation methodology for data association is divided into two tasks: ground truth development and an evaluation process, as discussed below.

3.1.5.6.2.1 Ground Truth Development

Development of the ground truth is a key step for evaluating the performance of any data association algorithm. The ground truth is typically prepared by one or more human analysts and it represents the answer key to the data association solution, against which the association algorithm is graded. The *soft ground truth* contains a list of unique entities, events and relationships with a unique identifier (UID) assigned to each of them; and another list containing observations of the unique entities, events and relationships (with respective UIDs) in various soft messages. The analyst also records the pedigree information related to each of entity, which represents the exact location and number of characters in the textual description of that entity in a

particular text message. The *hard ground truth* contains similar lists of unique and observed entities and events, present in each of the hard data sources, with cross-modality UIDs carried forward from soft data ground truthing.

3.1.5.6.2.2 Evaluation Process

As mentioned before, the performance of data association is assessed at two levels. For assessing the cumulative system performance (the "system perspective") at the data association process, the CDG is compared with the ground truth and three types of entity pairs are counted: (a) *correctly associated*; (b) *incorrectly associated*; and (c) *incorrectly not associated*. These counts are obtained by programmatically comparing the pedigree records of the nodes in the CDG with those of the entity observations in the ground truth. After obtaining these counts, we quantify the performance of the data association, using Precision, Recall, and F-score, which are defined below.

- *Precision:* Ratio of correctly associated entity pairs to the total number of associated entity pairs (i.e.$\frac{\mathrm{a}}{\mathrm{a+b}}$).
- *Recall*: Ratio of correctly associated entity pairs to the total number of correctly associated and incorrectly not associated entity pairs (i.e.$\frac{\mathrm{a}}{\mathrm{a+c}}$).
- *F-score*: Harmonic mean of the Precision and Recall values i.e. $\frac{2\times\mathrm{precision}\times\mathrm{recall}}{\mathrm{precision}+\mathrm{recall}}$.

The higher values of these metrics typically indicate greater accuracy. Since maximizing Precision and Recall are competing objectives, our focus is on maximizing the F-score. For this purpose, we trained a logistic regression model on the feature scores for a separate training dataset. The training algorithm calculates the optimal values of the feature weights used in similarity score calculation, with an objective of maximizing the F-score. For a more in depth description of the scoring and evaluation processes, readers are directed to [E.12].

Assessing the "data association perspective" performance of data association is not so straightforward, because any imprecision in the upstream processes could influence the data association results. Two examples of such imprecisions are: incorrect or missing entity typing and incorrect or missing within message co-referencing. To assess the standalone performance of data association (the "data association perspective"), we need to identify and disregard imprecisions in the data association solution stemming from upstream processes. To this end, we will explain the *type restricted evaluation* method, which helps in isolating association performance on "correct" input data (see [E.12] for a detailed explanation). In this method, we identify the entity pairs which are incorrectly associated or incorrectly not associated due to NLU errors; and disregard them from the counts (b) and (c) mentioned above. To prevent an unfair inflation of the Precision and Recall, we also identify the correct associations which overcame the NLU errors, and disregard them from the count (a). Using these counts, we can calculate the "data association perspective" Precision, Recall and F-score for data association, which are likely higher than their "system perspective" counterparts.

Note that our current association perspective evaluation strategy does not support the nullification of the effects of within message co-referencing errors. However, modeling data association as clique partitioning problem (CPP) helps recover some of the missing within message co-references and improves the F-score (as seen in Table 20).

3.1.5.6.3 Testing

We tested our evaluation strategy on the three data association formulations and corresponding algorithms: sequential Lagrangian heuristic for GA^N, Map/Reduce Lagrangian heuristic for MDADC and streaming entity resolution algorithm for CPP (see [E.28]-[E.30]). The procedures were coded in Java and executed on Intel Core 2 Duo processor, with 3 GHz clock speed and 4GB RAM. We have used a sample vignette message set of SYNCOIN as the input data set, which contains 114 soft messages and 13 hard messages. The statistics related to the evaluation engine are presented in Table 19, and the computational results for the data association algorithms are presented in Table 20.

Overall 46,030 pairs of pedigree records were compared during the evaluation process, of which 1,302 are within-message and 44,728 are between-message. We see that the association perspective evaluation (the lower row performance metrics within Table 20) results in higher Precision, Recall, and F-score, as expected.

The sequential Lagrangian procedure for GA^N formulation takes the second longest time to solve because of the complexity of the model. The Map/Reduce Lagrangian procedure for MDADC is quite fast, as a result of parallelization. Thus, for large graphs, the sequential Lagrangian heuristic for GA^N will prove to be a bottleneck. On the other hand, MDADC formulation solved using Map/Reduce can potentially provide a quick and accurate solution and it is easily scalable for larger graphs. The cumulative time required for Streaming Entity Resolution algorithm, is the largest; however it takes only 10 seconds per graph update. Streaming resolution also helps recover the missing within-message associations, improving the Recall of the system perspective evaluation.

Table 19: Evaluation statistics for sequential GA^N

Evaluation Mode	Correctly Associated	Incorrectly Associated	Incorrectly Not Associated
System Perspective	30,563	2,708	12,759
Association Perspective	29,349	2,382	8,836

Table 20: System (upper row) and association perspective (lower row) association performance by algorithm

<table>
<tr><th>No.</th><th>Procedure</th><th>Precision</th><th>Recall</th><th>F-Score</th><th>Compute Time (s)</th></tr>
<tr><td rowspan="2">1</td><td rowspan="2">GA^N (Sequential)</td><td>0.918</td><td>0.705</td><td>0.798</td><td rowspan="2">794</td></tr>
<tr><td>0.925</td><td>0.768</td><td>0.839</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">MDADC (MR)</td><td>0.932</td><td>0.708</td><td>0.805</td><td rowspan="2">64</td></tr>
<tr><td>0.938</td><td>0.772</td><td>0.847</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">CPP (Streaming)</td><td>0.909</td><td>0.730</td><td>0.810</td><td rowspan="2">1,312
(10 s/graph update)</td></tr>
<tr><td>0.915</td><td>0.796</td><td>0.851</td></tr>
</table>

3.1.5.7 Sensemaking via Graph Analytic Processes

The situation assessment processes within the SUT utilize as input the cumulative associated data graph formed by the data association process. The graph analytic processes for situation

assessment within our SUT are representative of just one analytic strategy for a hard+soft information fusion system, but they can be examined to illustrate some of the complexities of the broader evaluation issues for automated tools designed to aid sensemaking.

There are two major aspects for assessing a toolkit of automated methods to support a human-based sensemaking process: the performance of the algorithms in forming automated situational *assessments* (algorithmically-formed hypotheses), and the (possibly-separate) ability of these algorithms to aid in the formation of human-based situational *awareness*. While an automated algorithm (e.g., graph matching) may be *efficient* in *assessing* matches to specified situations of interest, this technique in itself may not be *effective* in supporting domain-wide *awareness***.** This is in part because of the underlying discovery/learning-based approach to sensemaking and the limitations of deep knowledge in modern problem domains such as counterinsurgency (COIN). In complex and dynamic problem environments like these, even the best assessment-supporting technologies are of limited capability today and many produce what we will call "situational fragments," *partial* hypotheses representing situational substructures as patterns. Situational awareness at a more complete level is the result of a dynamic interaction with the assessment tools, possibly using other technology to connect these "fragments" (as the human is trying to do) and human judgment in a kind of mixed-initiative operation. The evaluation focus of the graph-analytic tools in our SUT is on measuring the situation *assessment* capabilities, with the evaluation of effectiveness in developing situational awareness left for future work (see Section 3.1.5.8).

Three graph analytic processes within our SUT have been previously evaluated: a link analysis tool, social networking tool and stochastic graph matching tool. The algorithmic computational efficiency, specifically with a focus on data size scalability, of the link analysis algorithm is described in [E.31]. The evaluation of the social network tool for social network extraction and high value individual (HVI) identification is described in [E.1]. Finally, the evaluation of the stochastic graph matching tool to efficiently identify situations of interest within the cumulative associated data is presented in [E.32].

3.1.5.8 Discussion and Future Work

The example SUT and evaluation point process and system level performance metrics form the basis for error audit trail analysis. Through the utilization of this error audit trail numerous questions can be answered within the test space as described in Section 3.1.5.2, for example: What is the value of hard+soft fusion (versus hard only or soft only) toward some system level objective? While the answer to this and other experimental questions is ultimately the goal of this approach in systemic testing, we are currently still completing the training phase of this effort. In addition to the assessment of the evaluation questions listed in Section 3.1.5.2 on an independent test data set, other issues in the evaluation of hard+soft information systems remain as future work. Additional questions which will be assessed as future work include: how does one assess generality of methods on independent training and test data[4]? What are the challenges of testing in a streaming environment and how are performance metrics in tune with the dynamic user requirements within these environments? What are the dimensions of scalability which must

[4] We recognize there is some existing literature in the area of quantifying characteristics of a textual corpus via: statistical vocabulary analysis (lexicometry [E.33]), textural complexity (textometry [E.34]) and linguistic style (stylometry [E.35]) among other approaches. The investigation of these measures as an argument for framework generality remains as future work.

be considered both in input data and decision dissemination? What is the relationship between situational awareness and the resulting actions taken?

3.1.5.9 Conclusions

This paper presented a metric-based test and evaluation (T&E) framework for the assessment of a hard+soft fusion system. Issues in the definition of a System Under Test (SUT) and evaluation points in an active Research and Development program were discussed. An example SUT from the MURI Network-based Hard+Soft Information Fusion project is considered, with evaluation metrics at both the "process" and "system" level for each evaluation point provided. The future use of the evaluation framework in assessing design alternatives and incremental research and development efforts is also provided.

3.1.5.10 References

[E.1] K. Date, G. A. Gross, S. Khopkar, R. Nagi, K. Sambhoos. 2013. "Data association and graph analytical processing of hard and soft intelligence data," Proceedings of the 16th International Conference on Information Fusion (Fusion 2013), Istanbul, Turkey, 09-12 July 2013.

[E.2] C. L. Hornbaker II, "Tactical Fusion Centers: Restructuring Intelligence for Counterinsurgency," M.A. Thesis, (2012).

[E.3] J. Llinas "Information Fusion for Natural and Man-Made Disasters." Proc. of The 5th International Conference on Information Fusion, Annapolis, MD, 570-77, (2002).

[E.4] D. L. Hall and J. M. Jordan. "Information Fusion for Civilians: The Prospects of Mega-Collaboration." *Human-centered Information Fusion*. Boston: Artech House, 2010. 211-26.

[E.5] P. V. Puttan, J. N. Kok and A. Gupta, "Data Fusion through Statistical Matching," (2002).

[E.6] "Information fusion within the retail sector," letter, <http://archive.his.se/PageFiles/7158/ICA_scanned.pdf>, (2004).

[E.7] "Unified Research on Network-based Hard/Soft Information Fusion", Multidisciplinary University Research Initiative (MURI) grant (Number W911NF-09-1-0392) by the US Army Research Office (ARO) to University at Buffalo (SUNY) and partner institutions.

[E.8] D. L. Hall, J. Graham, L. D. More and J. C. Rimland. "Test and Evaluation of Soft/Hard Information Fusion Systems: A Test Environment, Methodology and Initial Data Sets." Proc. of The 13th International Conference on Information Fusion, Edinburgh, Scotland (2010).

[E.9] E. Blasch, P. Valin, and E. Bosse. "Measures of Effectiveness for High-Level Fusion." Proc. of The 13th International Conference on Information Fusion, Edinburgh, Scotland, (2010).

[E.10] M. R. Endsley, "Measurement of Situation Awareness in Dynamic Systems." Human Factors: The Journal of the Human Factors and Ergonomics Society 37.1 (1995): 65-84.

[E.11] J. J. Salerno, E. P. Blasch, M. Hinman, and D. M. Boulware. "Evaluating Algorithmic Techniques in Supporting Situation Awareness." Multisensor, Multisource Information Fusion: ARCHITECTURES, ALGORITHMS, AND APPLICATIONS. Ed. B. V. Dasarathy. S.l.: Int'L Soc For Optical, 2005. 96-104.

[E.12] K. Date, G. A. Gross, R. Nagi. "Test and Evaluation of Data Association Algorithms in Hard+Soft Data Fusion," Proc. of the 17th International Conference on Information Fusion, Salamanca, Spain, (2014).

[E.13] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discrimi- natively trained part based models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pp. 1627–1645, 2010.

[E.14] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," International Journal of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010.

[E.15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proceedings of IEEE International Conference on Computer Vision, vol. 2, pp. 886–893, 2005.

[E.16] H. Pirsiavash, D. Ramanan, C. Fowlkes. "Globally-Optimal Greedy Algorithms for Tracking a Variable Number of Objects" Computer Vision and Pattern Recognition(CVPR) Colorado Springs, Colorado, June 2011.

[E.17] M. Prentice, M. Kandefer, and S. C. Shapiro, "Tractor: A framework for soft information fusion," in Proceedings of the 13th International Conference on Information Fusion (Fusion2010), 2010, p. Th3.2.2.

[E.18] S. C. Shapiro and D. R. Schlegel, ``Natural language understanding for soft information fusion," in Proceedings of the 16th International Conference on Information Fusion (Fusion 2013), 2013, 9 pages, unpaginated.

[E.19] S. C. Shapiro and W. J. Rapaport, "The SNePS family," Computers & Mathematics with Applications, vol. 23, no. 2-5, pp. 243-275, January-March 1992, reprinted in F. Lehmann, Ed., Semantic Networks in Artificial Intelligence, Oxford: Pergamon Press, 1992, pp. 243-275.

[E.20] S. C. Shapiro, "An introduction to SNePS 3," in Conceptual Structures: Logical, Linguistic, and Computational Issues, Lecture Notes in Artificial Intelligence, B. Ganter and G. W. Mineau, Eds. Berlin: Springer-Verlag, 2000, vol. 1867, pp. 510-524.

[E.21] D. R. Schlegel and S. C. Shapiro, "Visually interacting with a knowledge base using frames, logic, and propositional graphs," in Graph Structures for Knowledge

Representation and Reasoning, Lecture Notes in Artificial Intelligence, M. Croitoru, S. Rudolph, N. Wilson, J. Howse, and O. Corby, Eds. Berlin: Springer-Verlag, 2012, vol. 7205, pp. 188-207.

[E.22] M. Palmer, VerbNet: A Class-Based Verb Lexicon, University of Colorado Boulder. http://verbs.colorado.edu/~mpalmer/projects/verbnet.html. Last accessed 27 February 2014.

[E.23] Princeton University "About WordNet." WordNet. Princeton University. 2010. http://wordnet.princeton.edu. Last accessed 27 February 2014.

[E.24] National Geospatial-Intelligence Agency, NGA GEOnet Names Server, http://earth-info.nga.mil/gns/html/. Last accessed 27 February 2014.

[E.25] C. van Rijsbergen, Information Retrieval, Second Edition, London: Butterworths, 1979.

[E.26] M. Jenkins, G. Gross, A. Bisantz, R. Nagi, “Towards Context Aware Data Fusion: Modeling and Integration of Situationally Qualified Human Observations into a Fusion Process for Intelligence Analysis,” Journal of Information Fusion, (Accepted June 2013).

[E.27] M.P. Jenkins, G. A. Gross., A. M. Bisantz, and R. Nagi, “Towards context-aware hard/soft information fusion: Incorporation of situationally qualified human observations into a fusion process for intelligence analysis.” In Proc. of the 2011 IEEE First International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), Feb. 22-24, Miami Beach, FL, pp. 74-81.

[E.28] G. Tauer, R. Nagi, and M. Sudit, “The graph association problem: mathematical models and a Lagrangian heuristic”. Naval Research Logistics, vol. 60, pp. 251-268, April 2013.

[E.29] G. Tauer, K. Date, R. Nagi, and M. Sudit, “An incremental graph-partitioning algorithm for entity resolution,” Under Revision. To be submitted to Transactions on Knowledge Discovery from Data.

[E.30] G. Tauer, and R. Nagi, “A Map-Reduce Lagrangian heuristic for multidimensional assignment problems with decomposable costs,” Parallel Computing, 39(11), pp. 653-658, November 2013.

[E.31] G.A. Gross, Graph Analytic Techniques in Uncertain Environments: Graph Matching and Link Analysis, Dissertation, University at Buffalo, 2013.

[E.32] G. A. Gross, R. Nagi, and K. Sambhoos. "A Fuzzy Graph Matching Approach in Intelligence Analysis and Maintenance of Continuous Situational Awareness." Information Fusion 18 (2014): 43-61.

[E.33] Z. Harris, “Mathematical Structure of Language.” John Wiley, New York, 1968.

[E.34] I. Aydin, and E. Seker, “Textometry: A Method for Numerical Representation of a Text,” International Journal of Humanities and Social Science Vol. 2 No. 23; December 2012.

[E.35] M. Brennan, et al., “Adversarial Stylometry: Circumventing Authorship Recognition to Preserve Privacy and Anonymity,” ACM Transactions on Information and System Security, Vol. 15, No. 3, Article 12, November 2012.

3.1.6 Social Network Analysis and High Value Individual Identification Evaluation

In order to evaluate the utility of a toolset for automatic processing of hard+soft data, one may be interested in revealing the relationships between the actors reported on in the data messages/signals. The objective of this section is to explore the tool’s potential for automatic identification of the person nodes of interest in the data and finding the most influential nodes, based on their structural positions in the relationship network, i.e., their social network; such influential individuals henceforth will be referred to as High Value Individuals (HVIs). This section explains our methodologies for: social network graph extraction from a Cumulative Associated Data Graph and use of Social Network Analysis (SNA) techniques to recognize HVIs [F.1], [F.2], [F.3]. The comparison of HVI identification results, obtained using the CDG extracted social network (CDGSN) versus the ground-truth social network (GTSN), allows one to assess the quality of the toolset.

3.1.6.1 Social Network Extraction and High Value Individual Modules

Extracting a social network, when the data reflecting direct underlying relationships between actors is not available, requires inference tools or data mining techniques [F.6], [F.7], [F.8]. Moreover, the comparisons in detecting HVIs must be done using multiple metrics, since different metrics capture different properties of actors’ structural positions in a social network [F.9]. The experiments with those different metrics are conducted using supervised learning, with multiple processing modules involved in the network extraction (see Figure 26).

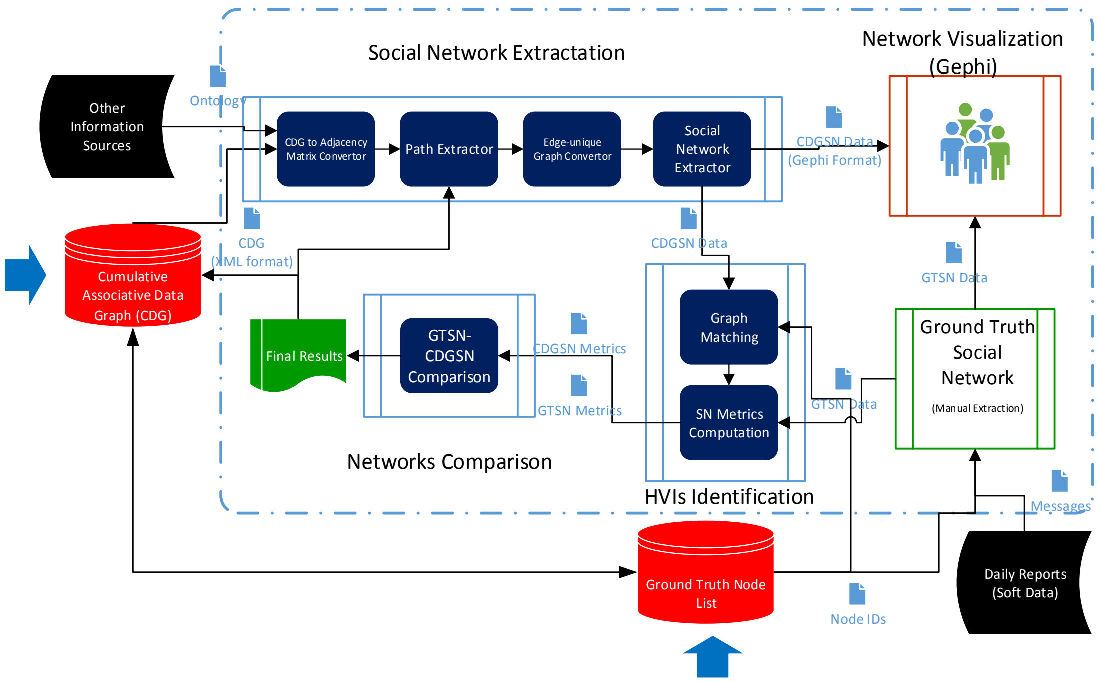


Figure 26: Social Network Extraction and High Value Individual Identification Modules

3.1.6.1.1 Cumulative Data Graph Social Network Extraction

The main idea employed in extracting a social network from CDG – Cumulative Associated Data Graph Social Network (CDGSN) - lies in traversing feasible paths between all the pairs of person-type nodes. The first step is to convert CDG XML formatted file to an adjacency matrix of all the nodes in CDG. In extracting the feasible paths for each pair of person-type nodes, the acceptance constraints are imposed to ensure that (1) no acceptable path contains a person node, and (2) the length of an acceptable path does not exceed a pre-set threshold (T). A DFS strategy is implemented where the time complexity for handling all the person node pairs is $O(n^2)$.

Weights are assigned to the formed edges between two people to incorporate the effective proximity (distance) in the realized network. An edge between a pair of people nodes who are two hops away should weigh less – in a distance sense – than an edge between a pair of people nodes who are three hops away from each other. Also, if multiple paths of lengths smaller than the hops threshold (T) exist between a pair of nodes, then the weight of the edge between those nodes should be smaller. To take into consideration these two effects, an edge weight is calculated by the following formula [F.21].

Let,

$w(i, j)$ be the weight of an edge between node i and node j.

P_{ij} be a set of all paths between node i and node j.

p_{ij} be a single path such that $p_{ij} \in P_{ij}$.

$h(p_{ij})$ be the number of hops in path p_{ij} .

T be the hops threshold.

Then,

$$w(i,j) = \frac{1}{\sum_{p_{ij} \in P_{ij}} \frac{1}{h(p_{ij})}}, \text{ such that } h(p_{ij}) \le T .$$

The inverse of this distance weight is used in identifying relationship strength (i.e., $1 - w(i,j)$).

3.1.6.1.2 High Value Individual Identification

Several methods have been proposed to identify the most influential nodes in a network [F.9], [F.10], [F.11]; the results obtained in this test and evaluation study are based on well-accepted social network analysis (SNA) measures of centrality. Note that the term "High Value" may have a particular meaning depending on a particular application.

In the SNA literature, centrality metrics such as betweenness, closeness, degree, etc., are interpreted as the prominence of actors embedded into a social network neighborhood [F.4][F.5][F.9][F.13]. Based on the desired prospective use of identified HVIs, the betweenness, closeness or degree based centrality definitions can be adopted for their identification, depending on the presumed nature of information/item exchange between the network actors.

Among these centrality metrics, the degree centrality can be viewed as a trivial metric (which reflects the number of direct connections a node has), while both the betweenness and closeness centralities are based on shortest paths connecting node pairs and measure the average distances from the nodes to their peers [F.12]. There exist many algorithms for calculating path-based centrality values. One of the most efficient algorithms, proposed by Brandes [F.12], is implemented for this study.

Prior to the identification of HVIs, a graph matching problem needs to be solved to match the nodes within the CDGSN and GTSN. Weights indicating the strength of a match between a CDGSN node (identified via the MURI processes of natural language processing, physical sensor processing and data association) and GTSN are calculated as the degree of overlap between pedigree records contained within the CDGSN node and ground truth. The methodology for this calculation is analogous to the data association evaluation technique detailed in Section 3.2. These weights form an assignment matrix between all nodes in the CDGSN and GTSN. The matching problem is thus formulated as a linear assignment problem, attempting to maximize the degree of overlap of the CDGSN-GTSN mapping.

3.1.6.1.3 Ground Truth Social Network Extraction

For each given dataset, a corresponding Ground Truth Social Network (GTSN) is extracted by social network experts. Reading and comprehending all the data messages, the experts identify all the actors involved, and come to a consensus on the implied edges (relationships) even if they are not directly mentioned in the messages or require information fusion across multiple messages.

Controversy may arise in certain cases, since two experts may or may not agree on the existence of a particular link. In order to resolve such issues, a set of ground rules is proposed by the GTSN extractor team. The resulting GTSNs play an important role in working with both training and test sets in the learning process.

3.1.6.1.4 Network Visualization

Apart from metrics calculations, automatic visualization helps an analyst to evaluate and compare HVI's positions. In addition, one can study the cohesiveness of the underlying networks and concentrate more on the links of interest.

To visualize the networks, an open-source software Gephi [F.14] is utilized. Gephi can handle directed, undirected, weighted, unweighted, attributed, static and dynamic graphs. The added functionality (introduced specifically for the needs of this project) allows one to seamlessly integrate and display network nodes and edges together with the message-based pedigree data.

3.1.6.1.5 Comparison between CDGSN and GTSN

Given two networks A and B, a simple way to assess their similarity is to count the number of changes that one has to do to transform one graph into the other (this measure is known as graph edit-distance [F.15]). Various edit operations have been introduced to date, including edge rotation, edge addition and subtraction, vertex addition and subtraction (if the networks do not have the same number of nodes), etc.; note that it is not obvious how to weigh these changes against one another [F.15].

There is a wide array of other methods proposed in the existing literature and exploited to measure the similarities between two given networks. Spectral analysis is used to approximate the graph edit-distance by the difference in the spectrum of eigenvalues between Laplacians of the adjacency matrices [F.16]. Other related research introduces p* models (now widely known as Exponential Random Graph Models [F.17]), graph kernels [F.18] and motif analysis [F.19]. The p* models and motif analysis are based on the presence of small subgraphs in the compared networks [F.17]. Graph kernels map graph features to points in high dimensional inner product spaces [F.15].

However, given the objective of HVI's identification as the common one in the intelligence domain, the HVI-based measures, e.g., rankings, can be directly used to evaluate the quality of the extracted network. With the ranking of aforementioned centrality metrics used to identify HVI's, Kendall Tau distance [F.20] can be accordingly utilized to compare the quality of CDGSNs relative to GTSNs.

Indeed, the Kendall Tau is a rank correlation coefficient, i.e., the statistic used to evaluate the association between two measured quantities [F.20] (e.g., betweenness of nodes in two different networks),

$$\tau = \frac{N_c - N_D}{\binom{N}{2}},$$

where N_c and N_D are the counts of concordance and discordance pairs, respectively, and $\binom{N}{2}$ is the total number of pairs. Any pair of observations (x_i, y_i) and (x_j, y_j) is termed concordant if the ranks for both elements are the same. In other words, if both $x_i > x_j$ and $y_i > y_j$ or if both $x_i < x_j$ and $y_i < y_j$. On the other hand, any pair of observations is termed discordant, if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant.

3.1.6.2 Test and Evaluation

In this part, the design of a metric for evaluating systemic error trail analysis and parametric optimization of the social network extraction and HVIs identification is described.

3.1.6.2.1 Evaluation Objectives

As explained above, the main utility of an extracted social network is assumed to lie in distinguishing HVIs. The main expected functionality of the extracted social network is to correctly pinpoint HVIs for any given datasets; therefore, the objective is to minimize HVI miss-identification probability, or the HVI ranking list divergence from that based on the ground truth.

3.1.6.2.2 Evaluation Metrics

The metric utilized for minimizing the number of the HVIs misplaced in the ranked list needs to allow for the comparison of ranks of the pre-selected centrality metrics for each node in both GTSN and CDGSN. It should be noted that CDGSN and GTSN are not necessarily expected to have the same number of nodes due to upstream processing errors. However, most or all ground truth nodes are expected to be present in both networks; in this case, the node ranks by centrality metric values can be compared using Kendall Tau distance.

Let B_i^C and B_i^Grepresent the betweenness values of node i in CDGSN and GTSN, respectively. Similarly, let C_i^C and C_i^G denote the closeness of node i in CDGSN and GTSN, and D_i^C and D_i^G denote degree of node i in CDGSN and GTSN ($\forall\, i \in\ \Omega$ where Ω is a set of common nodes in both CDGSN and GTSN). Considering the pair (B_i^C,B_i^G), the betweenness-based Kendall Tau (τ_B) is defined as follows:

$$\tau_B = \frac{|\Omega_{Co}^B| - |\Omega_{Di}^B|}{\binom{|\Omega|}{2}},$$

where $\Omega_{Co}^B$and Ω_{Di}^B denote sets of concordance and discordance pairs for betweenness, respectively and $|\Omega|$ is Ωcardinality. Similarly formulae, the closeness-based Kendall Tau (τ_C) and degree-based Kendall Tau (τ_D) for (C_i^C,C_i^G) and (D_i^C, D_i^G), respectively, are:

$$\tau_C = \frac{|\Omega_{Co}^C| - |\Omega_{Di}^C|}{\binom{|\Omega|}{2}} \text{ and } \tau_D = \frac{|\Omega_{Co}^D| - |\Omega_{Di}^D|}{\binom{|\Omega|}{2}}.$$

Again, Ω_{Co}^C, Ω_{Di}^C, Ω_{Co}^D and Ω_{Di}^D denote the sets of concordance and discordance pairs for closeness ands degree, respectively.

3.1.6.2.3 Training Methodology

The training process for detecting HVIs requires one parameter to be tuned; the parameter is called the hop threshold, $T\ \geq 1\ \&\ \in \mathbb{Z}$, used in the path extraction process. The goal is to find T^* as

$$T^* = \underset{T}{\operatorname{argmin}}(w_B\tau_B + w_C\tau_C + w_D\tau_D),$$

where w_B, w_C and w_D are the weights for τ_B, τ_C and τ_D, respectively. In the simplest case, $w_B = w_C = w_D = 1$.

3.1.6.3 Training and Test Data Set Results

As part of the larger test and evaluation experiment, this work is currently in progress.

3.1.6.3.1 CDGSN and GTSN visualization

An example visualization tool output example based on the training dataset depicting both GTSN and pruned version of CDGSN is given in Figure 27.

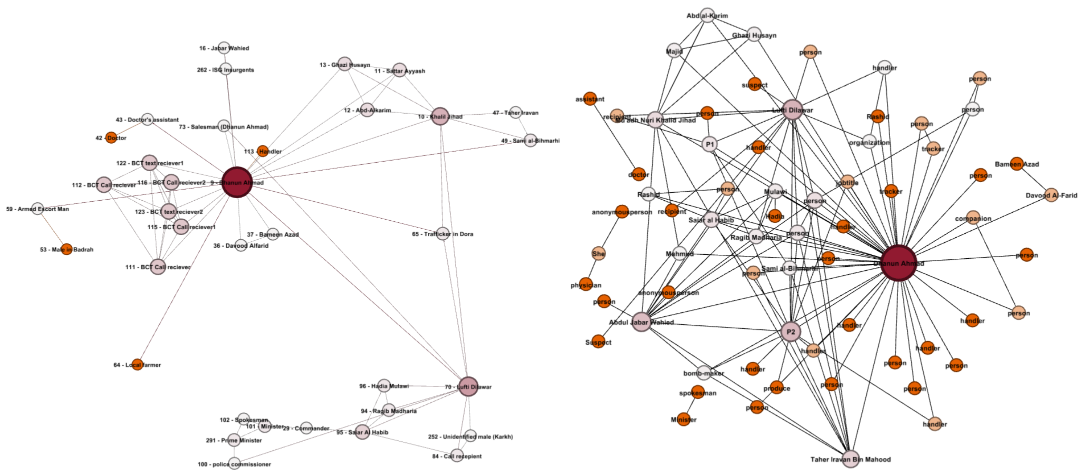


Figure 27: GTSN (left) and CDGSN (right)

3.1.6.3.2 HVIs in the training dataset

Based on the training dataset, the following preliminary results show the ranks of the HVIs based on three centralities metrics for top five HVIs ($T = 3$).

Table 21: Results for HVI ranking

Rank	Betweenness Rank in GT	Betweenness Rank in CDG	Closeness Rank in GT	Closeness Rank in CDG	Degree Rank in GT	Degree Rank in CDG
1	9	95	9	9	9	9
2	70	10	36	16	70	70
3	65	9	37	70	10	10
4	10	70	65	36	95	95
5	95	11	43	43	11	13

3.1.6.3.3 Kendall Tau Distance results for CDGSN and GTSN

Based on the training dataset, the following table shows $\tau_B + \tau_C + \tau_D$ for different values of T.

Table 22: Kendal tau values for training dataset

Path length thresholds (T)	Betweenness-based Kendall Tau (τ_B)	Close-based Kendall Tau (τ_C)	Degree-based Kendall Tau (τ_D)	Sum of Kendall Tau
1	0.323	0.418	0.513	1.254
2	0.316	0.435	0.5	1.251
3	0.303	0.392	0.515	1.210
4	0.375	0.414	0.479	1.268
5	0.331	0.443	0.51	1.284

In addition, the following figure illustrates the sum of centrality-based Kendall Tau for different values of path length threshold. The preliminary results show that the $T^* = 3$.

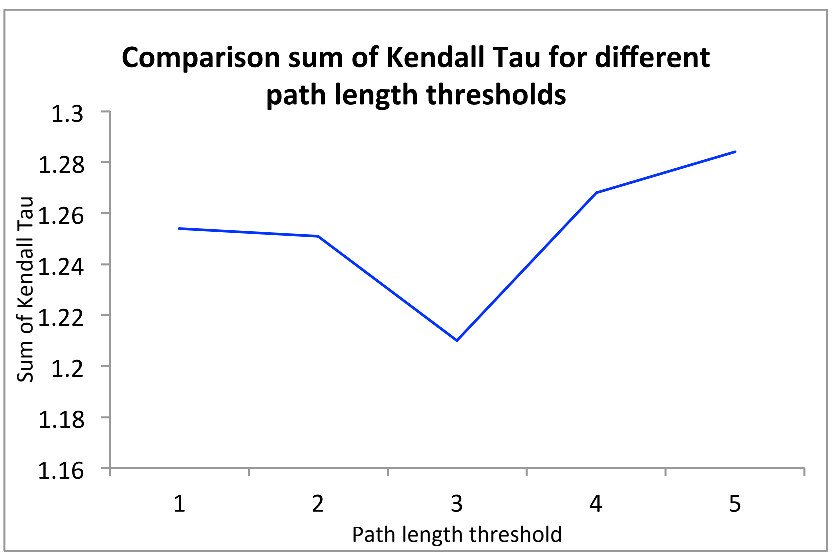


Figure 28: Kendal tau comparison

3.1.6.4 References

[F.1] Chen, D., Lü, L., Shang, M. S., Zhang, Y. C., & Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica a: Statistical mechanics and its applications*, *391*(4), 1777-1787.

[F.2] Suri, N. R., & Narahari, Y. (2008, May). Determining the top-k nodes in social networks using the shapley value. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3* (pp. 1509-1512). International Foundation for Autonomous Agents and Multiagent Systems.

[F.3] Kiss, C., & Bichler, M. (2008). Identification of influencers—measuring influence in customer networks. *Decision Support Systems*, *46*(1), 233-253.

[F.4] Tang, J., Musolesi, M., Mascolo, C., Latora, V., & Nicosia, V. (2010, April). Analysing information flows and key mediators through temporal centrality metrics. In *Proceedings of the 3rd Workshop on Social Network Systems* (p. 3). ACM.

[F.5] Varlamis, I., Eirinaki, M., & Louta, M. (2010, August). A study on social network metrics and their application in trust networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on*(pp. 168-175). IEEE.

[F.6] Jensen, D., & Neville, J. (2003, July). Data mining in social networks. In*Dynamic Social Network Modeling and Analysis: workshop summary and papers* (pp. 287-302). National Academies Press.

[F.7] Van der Aalst, W. M., & Song, M. (2004). Mining Social Networks: Uncovering interaction patterns in business processes. In *Business Process Management*(pp. 244-260). Springer Berlin Heidelberg.

[F.8] Thelwall, M., Wilkinson, D., & Uppal, S. (2010). Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the American Society for Information Science and Technology*, *61*(1), 190-199.

[F.9] Borgatti, S. P. (2006). Identifying sets of key players in a social network.*Computational & Mathematical Organization Theory*, *12*(1), 21-34.

[F.10] Shetty, J., & Adibi, J. (2005, August). Discovering important nodes through graph entropy the case of enron email database. In *Proceedings of the 3rd international workshop on Link discovery* (pp. 74-81). ACM.

[F.11] Arulselvan, A., Commander, C. W., Elefteriadou, L., & Pardalos, P. M. (2009). Detecting critical nodes in sparse graphs. *Computers & Operations Research*,*36*(7), 2193-2200.

[F.12] Brandes, U. (2001). A faster algorithm for betweenness centrality*. *Journal of Mathematical Sociology*, *25*(2), 163-177.

[F.13] Borgatti, S. P., & Halgin, D. S. (2011). Analyzing affiliation networks. *The Sage handbook of social network analysis*, 417-433.

[F.14] Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM*, *8*, 361-362.

[F.15] Macindoe, O., & Richards, W. (2010, August). Graph comparison using fine structure analysis. In *Social Computing (SocialCom), 2010 IEEE Second International Conference on* (pp. 193-200). IEEE.

[F.16] M. Peabody, (2002) "Finding groups of graphs in databases," Master's thesis, Drexel University, Philadelphia.

[F.17] Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social networks*, *29*(2), 173-191.
[F.18] Shervashidze, N., & Borgwardt, K. M. (2009). Fast subtree kernels on graphs. In *Advances in Neural Information Processing Systems* (pp. 1660-1668).
[F.19] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*,*298*(5594), 824-827.
[F.20] Fagin, R., Kumar, R., & Sivakumar, D. (2003). Comparing top k lists. *SIAM Journal on Discrete Mathematics*, *17*(1), 134-160.
[F.21] Date, K., Gross, G.A., Khopkar, S., Nagi, R., Sambhoos, K., "Data Association and Graph Analytical Processing of Hard and Soft Intelligence Data," Proc. of 2013 16th International Conference on Information Fusion (FUSION), Istanbul. p404-11.

3.1.7 Sensemaking and Argumentation

3.1.7.1 Sensemaking and Perspectives on Analytics

3.1.7.1.1 Introduction

The MURI program, in its efforts to develop a fully functional research prototype networked hard + soft data fusion capability, worked initially on all of the necessary front-end processing regarding ingestion and then the basic fusion process functions of Common Referencing and Association. In about the middle years, tasks were created to develop what we will call "focal" analytic tools such as a Link Analysis capability and a Social Network Analysis capability, along with the Graph-Matching tool that was imported from the ARL "STEF" program and then functionally expanded within the MURI. In the last one and a half years, the program has changed its focus onto two major functional areas: test and evaluation and integrated analytics. As regards prototyping of integrated analytics, the program largely focused on composite visualization schemes, an effort led by Penn State that has produced first versions of multi-pane visualization schemes that include inter-tool linking and agile visualization.

Although scope limitations and other factors prevented working toward a prototyping goal for an integrated analysis suite in this last year, it was decided to conduct a layered study addressing issues and design concepts for technology-based approaches that could support holistic human-machine Sensemaking and decision support. By holistic is meant support to the formation of synthesized, situational-level hypotheses that are aggregated in part from the "focal" hypotheses nominated by each of the tools mentioned above. Such technology would work in concert with the adaptive visualization designs created by Penn State.

At the highest level, we explored the design issues in forming a fully-connected and interdependent set of processes that include Information Fusion, Sensemaking, and Decision Support operations. Subsequently, we explored a particular thrust in integrated analytics based on a Story and Belief-based Argumentation scheme; this is addressed in Section 3.1.7.2 below.

3.1.7.1.2 Reexamining Information Fusion--Decision Making Inter-dependencies

In the Counterinsurgency (COIN) environment, typical of the modern complex military problem domain, current military doctrine envisions the COIN battlespace as a "mosaic" of localized situations, challenges, and possible solutions [G.1.4]. Developing estimates of this Battlespace and its situations requires an inferencing and estimation process that is a combination of deductive, inductive, and abductive processes that in turn exploit prior knowledge, multisource data, and experience. Further, the process is dynamic and iterative, adapting as hypotheses are discovered/nominated and verified, and the various hypotheses synthesized into a defendable and plausible whole, to achieve a "final" hypothesis that then can be used as a basis for decision-making. Collectively, these processes involve Information Fusion, Sensemaking, and then Decision-Making. It is important to understand that these processes are interdependent and that the design of any of these functional segments requires consideration of these interdependencies. Our first-layer study examined these issues and collected information and knowledge supportive of fully-integrated design of such processes. Two studies were conducted and these led to two conference papers [G.1.1], [G.1.2].

In [G.1.1], an integrated process model was nominated and discussed as regards the many factors that impact a systemic design approach, with a focus on functional and process interdependencies; that process model is shown below; interested readers are referred to [G.1.1] for details:

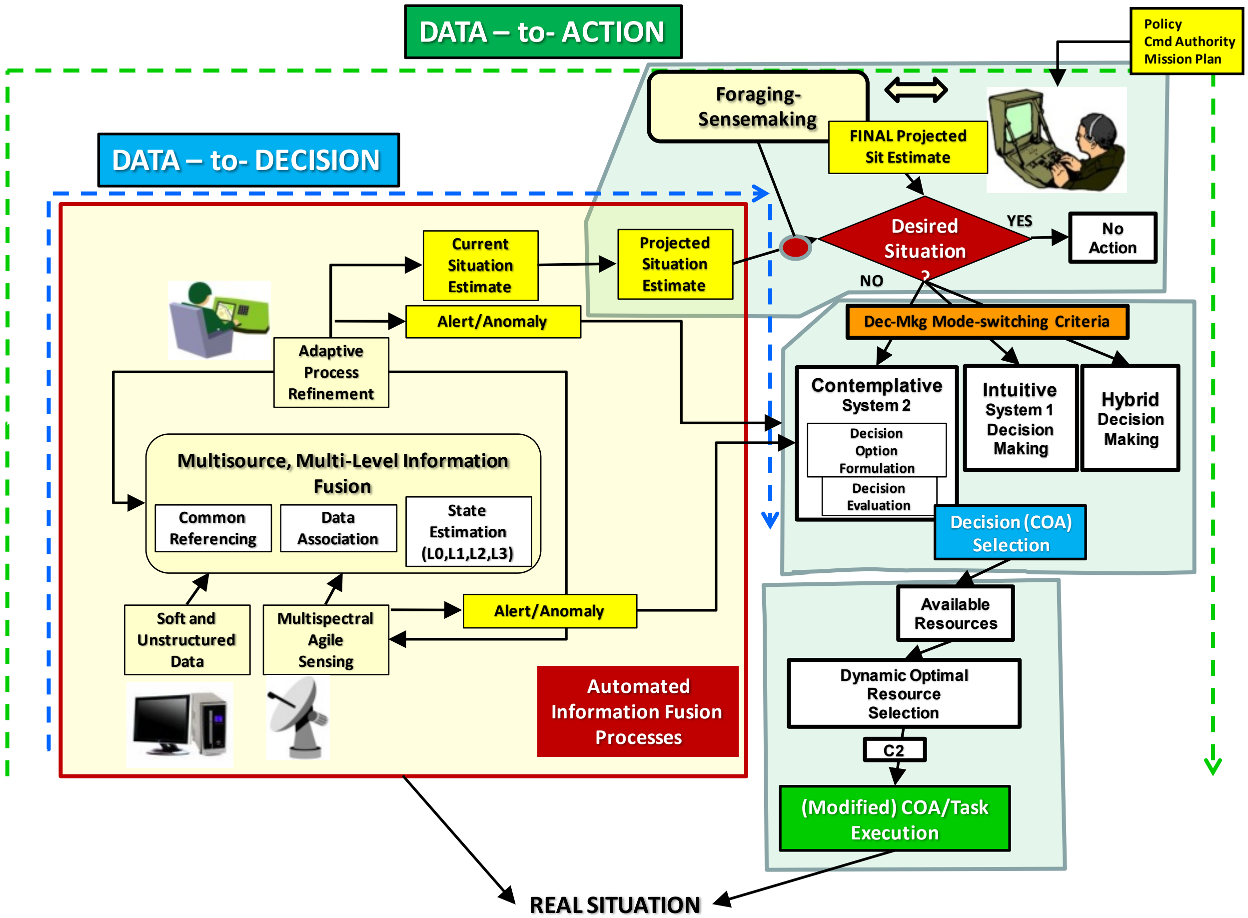


Figure 29: Interconnected information fusion – sensemaking – decision-making processes

In the second study [G.1.2], the focus was on exploring technologies that could provide a basis for synthesizing the hypotheses nominated by both the focal-type tools mentioned above as well as those nominated by a human analyst, so as to reduce the substantial cognitive workload on human analysts required to conduct a strictly-mentally-based integration to an integrated Battlespace/situational picture. It was this study that led to the nomination of argumentation-based methods that have been studied and described in Section 3.1.7.2 below. This literature survey examined papers in Law, in Critical Thinking, in Artificial Intelligence, and in Criminal Analysis domains in a fairly extensive study. In researching these methods, it was noted in [G.1.3], that van den Braack contended that an optimal approach to analysis would combine scenario or narrative-based techniques with argumentation techniques. A schematic representation of this combination is displayed in Figure 30, from [G.1.3] that illustrates that in the argument-based approach, arguments are constructed starting with a piece of evidence (see *evidential arguments* in Figure 30), and reasoning steps are performed to reach a conclusion based on this evidence, whereas in the story-based approach, stories about what might have happened are constructed in order to explain the evidence (the *scenario* in Figure 30). This design is prototyped in [3] with reasonably good results against a variety of criminal analysis use cases.

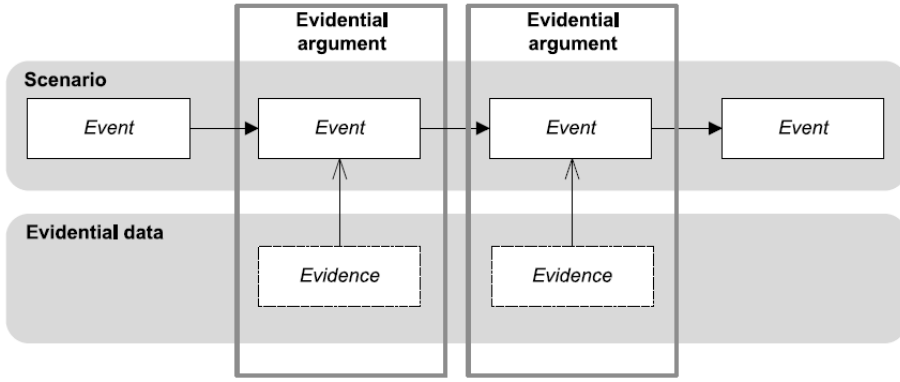


Figure 30: Notional Approach to Combining Story/Scenario-based and Argument-based Methods [G.1.3]

3.1.7.1.3 References

[G.1.1] Llinas, J., Reexamining Information Fusion--Decision Making Inter-dependencies, Presented at IEEE CogSIMA (Cognitive Situation Management) Conference, San Antonio, TX, Mar 2014.

[G.1.2] Llinas, J., A survey of automated methods for sensemaking support, Presented at the SPIE Next-Generation Analytics Conference (part of SPIE Defense and Security Conf), Baltimore, MD, April 2014.

[G.1.3] van den Braack, S., "Sensemaking software for crime analysis", Dutch Research School for Information and Knowledge Systems (SIKS), SIKS Dissertation Series No. 2010-12, 2010.

[G.1.4] Connable, B., "Embracing the Fog of War: ASSESSMENT AND METRICS IN COUNTERINSURGENCY", RAND Report, 2012 (http://www.rand.org/pubs/monographs/MG1086.html).

3.1.7.2 An Approach to Story and Belief-based Argumentation for Threat Assessment

3.1.7.2.1 Motivations and Modeling Framework

The *cumulative associated data graph*[5] resulting from the hard+soft data association process developed in years 1-5 does not directly provide abduction capability necessary for explaining the associated data. Pieces of information contained in the *cumulative associated data graph* represent possibly uncertain and unreliable pieces of evidence supporting alternative stories (hypotheses) about a real-world situation.

[5] Note – the cumulative associated data graph contains fused, potentially uncertain data, associated across the hard+soft data modalities.

The objective of this research is to develop a new technology-based approach to support synergistic human-machine dynamic abductive reasoning over the pieces of evidence providing the capability of aiding knowledge discovery for supporting an analyst in recognizing human activity, and detecting and identifying potential and imminent threats with higher confidence and reduced cognitive workload.

The developed approach takes into account the specific characteristics of the environment, in which threat detection and recognition has to take place, such as:

- A Noisy and uncertain highly dynamic environment with insufficient, and in many cases, non-existing *a priori* statistical information
- Large amounts of data and information, often uncertain, some irrelevant, contradictory, conflicting, and unreliable
- Resource and time constraints, high cost of error.
- An Open world, in which something unexpected or even unimaginable can happen.
- The lack of complete knowledge bases to support analysis/reasoning.

These environmental characteristics shape the modeling framework of the proposed processing, which comprises belief-based argumentation combining the Transferable Belief Model [TBM] with a story-based defeasible argumentation, any-time decision making approach. The modeling framework integrates current MURI capabilities to include the results of soft/hard data association, graph matching techniques and other analytic outputs on the cumulative associated data (e.g., social network analysis, link analysis, etc.).

The TBM [G.2.1] is a two-level model for representing the quantified belief held by an agent at a given time on a given frame of discernment. Quantified beliefs in hypotheses about an object or state of the environment are represented and combined at the *credal* level while decisions are made based on probabilities obtained from the combined belief by the *pignistic* transformation at the *pignistic level*. The following section expands on and clarifies these ideas.

Formally, let Θ be a set of atomic hypotheses about the state of the environment or an identity of an object: $\Theta = \{\theta_1,...,\theta_k\}$. Let 2^Θ denote the power set. A function m is called a basic belief assignment (bba) if $m: 2^\Theta \to [0,1], \quad \sum_{A\subseteq\Theta} m(A) = 1.$

In the majority of belief models (see, e.g. [G.2.8],[G.2.9],[G.2.10]) $m(\varnothing)$ (uncommitted belief) is defined as zero (which invokes a closed world assumption) while the TBM is the only belief model, in which uncommitted belief can be non-zero, allowing for an open-world framework.

If m_1 and m_2 are basic belief assignments defined on Θ, they can be combined at the *credal level* with TBM by conjunctive combination or an unnormalized Dempster's rule, defined as:

$$m^\Theta(A) = \sum_{B\cap D=A} m_1(B)m_2(D), \ \forall A \subseteq \Theta \qquad (1)$$

Normalization of the combination rule in the belief models is performed by redistributing $m(\varnothing)$ among other subsets of Θ to obtain $m(\varnothing)=0$ for the combination result. The most popular normalized rule is the normalized Dempster's rule of combination:

$$m_{norm}^{\Theta}(A)=\frac{1}{1-K}\sum_{B\cap D=A} m_1(B)m_2(D),\ \forall A\subseteq\Theta \tag{2}$$

$$K=\sum_{B\cap D=\varnothing} m_1(B)m_2(D). \tag{3}$$

K is usually called "conflict."

Decision making is carried out at the *pignistic level* by using *pignistic probability*:

$$BetP^{\Theta}(A)=\sum_{B\subseteq\Theta}\frac{|A\cap B|}{|B|}\frac{m^{\Theta}(B)}{1-m(\varnothing)},\quad \forall A\subseteq\Theta\ , \tag{4}$$

where $|A|$ is the number of elements of Θ in A.

The TBM is very appropriate for representing belief in the complex COIN environment. Beliefs represented in the TBM "do not ask for explicit underlying probability functions" [G.2.1],[G.2.2]. They are sub-additive, which permits for numerically expressing uncertainty and ignorance. The TBM combination rule also allows for incorporation of belief reliability. Moreover, the TBM works under the open world assumption, i.e. it does not assume that the set of hypotheses under consideration is exhaustive. These properties of the TBM have been successfully exploited, e.g., for tracking, target recognition, and situation assessment [G.2.2]-[G.2.6].

Anytime decision making models are designed to support time-critical decision making and actions. They offer a means to improve decision quality over time, which may be improved gradually for example as more observations are available [G.2.6]-[G.2.9]. It is important to notice that decision quality depends on the problem at hand and the problem context [G.2.10]. Utilization of an anytime decision model for threat recognition is dictated by the fact that dealing with threat requires timely decisions and swift actions. Waiting may result in unacceptable decision latency leading to significant damage and casualties. At the same time, the false alarms can result in the costly disruption of regular activities and the waste of valuable resources.

Argumentation is recognized in the literature (see, e.g. [G.2.11]-[G.2.16]) as a promising method for defeasible reasoning with vague, inconsistent, incomplete knowledge and has been used in multiple domains such as legal [G.2.17], education [G.2.18] and cooperative decision making [G.2.19] as a decision aid tool. It is based on the construction, combination and comparison of arguments for and against certain hypotheses. An argument-based framework for decision making allows for explicitly following the rational decision processes of agents, which explain and justify agent's preferences over alternative hypotheses. There have been multiple argumentation schemes developed with each of them having advantages and drawbacks as methods useful for supporting decisions for threat recognition.

3.1.7.2.2 Argumentation theories

The majority of argumentation-based methods utilize a deterministic formal logic and theorem proof and the notion of argument acceptance and attack, e.g. Dung's theory of argumentation [G.2.13]. These methods require:

- Definition of the component parts of an argument and their interaction.
- Identification of rules and protocols describing argumentation processes.
- Methods for distinguishing legitimate from invalid arguments.
- Determination of conditions under which further discussion is redundant.

A similar paradigm allowing for default reasoning and various non-monotonic logics, namely the assumption-based framework, was defined in [G.2.2],[G.2.21]. Assumption-based argumentation considers arguments not as atomic elements but as deductions of a conclusion based on a set of assumptions. Assumptions are defined as inference rules, which may represent causal information, argument schemes, or laws and regulations. In general, abstract argumentation is "a tool for analyzing particular argumentation systems and for developing a meta theory of such systems, and not as a formalism for directly representing argumentation-based inference" [G.2.12]. While there were several publications describing "instantiation" of this paradigm see, e.g. [G.2.12],[G.2.22],[G.2.23],[G.2.24], the logic-based automatic argumentation scheme has several drawbacks as applied to the problem of threat assessment. As an example, they perform reasoning based on information in their knowledge base, which requires constant update. They are characterized by high computational complexity and do not allow for explicit incorporation of uncertainty.

The problem of explicit incorporation of uncertainty was addressed in the Probabilistic Argumentation Systems (PAS) [G.2.25] and Belief-base Argumentation Systems (BAS) in [G.2.4],[G.2.5]. PAS combines symbolic logic with probability theory and is useful for reasoning as an uncertain environment extension of the Assumption-Based Argumentation. PAS is characterized by a knowledge base containing propositions and uncertain assumptions as well as *a priori* probabilities that assumptions are true. Arguments supporting (or refuting) certain hypotheses are the conjunction of propositions and assumptions for which hypotheses are true (or false). The support of each hypothesis is defined as the disjunction of all minimal arguments supporting a hypothesis. The BAS is a modification of the PAS, in which *a priori* probabilities are replaced by subjective beliefs dynamically assigned to uncertain assumptions. While these argumentation systems represent a welcome extension of formal argumentation theories they still require a knowledge base and are characterized by high computational complexity.

A different type of argumentation systems is an "argument assistant system," which, in contrast to an automatic reasoning tool based on logic and theorem proving represents "an argument assistant" guiding the user's production of arguments and managing the argumentation process [G.2.26]. Argument assistant systems are designed to [G.2.26]:

- keep track of the issues that are raised and the assumptions that are made,
- keep track of the reasons that have been adduced for and against a conclusion,
- evaluate the justification status of the statements made, and
- check whether the users of the system obey the pertaining rules of argument.

Such systems are designed for various domains such as legal, educational, cooperative decision making and implemented in different fashions such as template-based, story based, or a combination of story-based with abductive reasoning systems. There are multiple existing software packages implemented as argument assistants. Such systems are more useful for analyzing the stream of data and making sense of it than automatic systems since they do not only rely on knowledge bases , allowing analysts to be more proactive and perform real-time refinement of the set of arguments and argumentation schemes based on what they see. They can perform imaginative discovery of novel arguments by having in mind that there may be certain unexpected arguments in the open world. The analyst can frame queries, which give context to the search for, and prioritization of, relevant arguments and hypotheses. One of the most promising user support systems is Araucaria [G.2.27], an open source software, which provides a user visualization of the argument structure of a text, while assisting in the drafting of the argumentation structure of a text by allowing manually dragging text into a graph structure that represents the argumentation.

While most pure argument-based approaches to analyzing an information stream are deductive and support or reject a certain hypothesis, a system supporting an analyst in threat detection and evaluation requires an understanding of “what happened” based on the evidence (abduction). The most promising approach introduced in [G.2.28],[G.2.29] is a hybrid theory, a theory for best explanation where “causal stories are hypothesized to explain the evidence, after which these stories can be supported and attacked using evidential arguments. For example, arguments can be used to further support a story with evidence or to reason about the plausibility of a causal link in a story” [G.2.29]. We judge that this approach has potential as an extension of the MURI analysis tool suite, in which dynamically analyzed associated and fused soft and hard pieces of data and information can be considered as pieces of stories. The major drawback of the hybrid theory described in the literature is that it does not explicitly incorporate uncertainty associated with the information representing these pieces of stories and reliability of the source of these data and information.

3.1.7.2.3 Belief-based hybrid argumentation.

The approach proposed in this research was developed to overcome the drawback of existing argumentation models. It considers a variation of the hybrid story-based model, which combines *pro* and *contra* arguments built from uncertain transient information while seeing each piece of this information as an element of alternative stories (hypotheses based on “what might happen”), with the TBM allowing for assigning beliefs to each argument, combining these beliefs, and selecting a story (hypothesis) based on the highest pignistic probability. A top-level functional diagram of this model is presented in Figure 31.

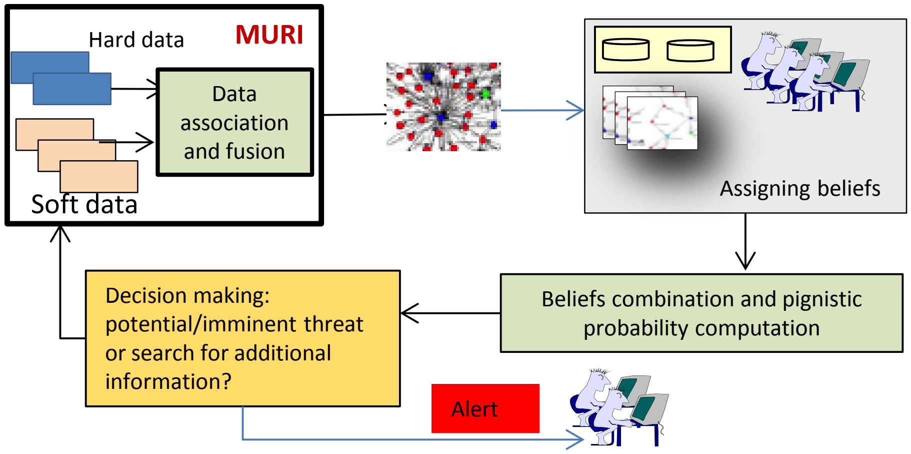


Figure 31: Belief-based hybrid argumentation (functional diagram)

We will discuss each processing module of this diagram below.

3.1.7.2.3.1 Building arguments and assigning beliefs.

This subsection describes the process of building arguments from the *cumulative associated data graph* resulting from association and fusion of the flow of soft and hard data. We consider three variants of the process shown in Figure 31, which differ by the level of human involvement in argumentation construction: a pure human-based process, a human-computer-based process (which includes automatic argument mining), and story-based argument construction. These three process variants are presented in Figure 32, Figure 33 and Figure 34, respectively.

During each time interval $[t_i, t_{i+1}]$, the cumulative associated data graph is utilized in designing arguments pro and contra threat hypotheses. Threat is defined as an integrated whole of three inter-related parts [G.2.28]:

- **Intentions**: plans or goals to be accomplished. These represent the psychological component of threats and can be deeply influenced by one's capabilities and opportunities. .
- **Capabilities (i.e., capacities)**: the kinds of objects (e.g., weapons), object attributes (e.g., projectile or explosive abilities) or behaviors (e.g., movements, perceptual abilities) that can inflict a certain level of harm, disruption or lethality on some target (as identified by one's intentions and made available by opportunities).
- **Opportunities**: the spatio-temporal states of affairs like access to a person or facility, abilities to know the adversary's plans (intentions). Opportunities makes it possible to actualize (i.e., carry out) one's intent given sufficient capabilities.

Since analysts are interested in recognizing not only imminent threat, which is characterized by the existence of all three threat components, but also potential threat characterized by existence of any two of these components, arguments supporting and refuting each of these components should be considered separately. This requires a sub-process of filtering the data graph into subgraphs with each of them containing information related to each of the three threat components.

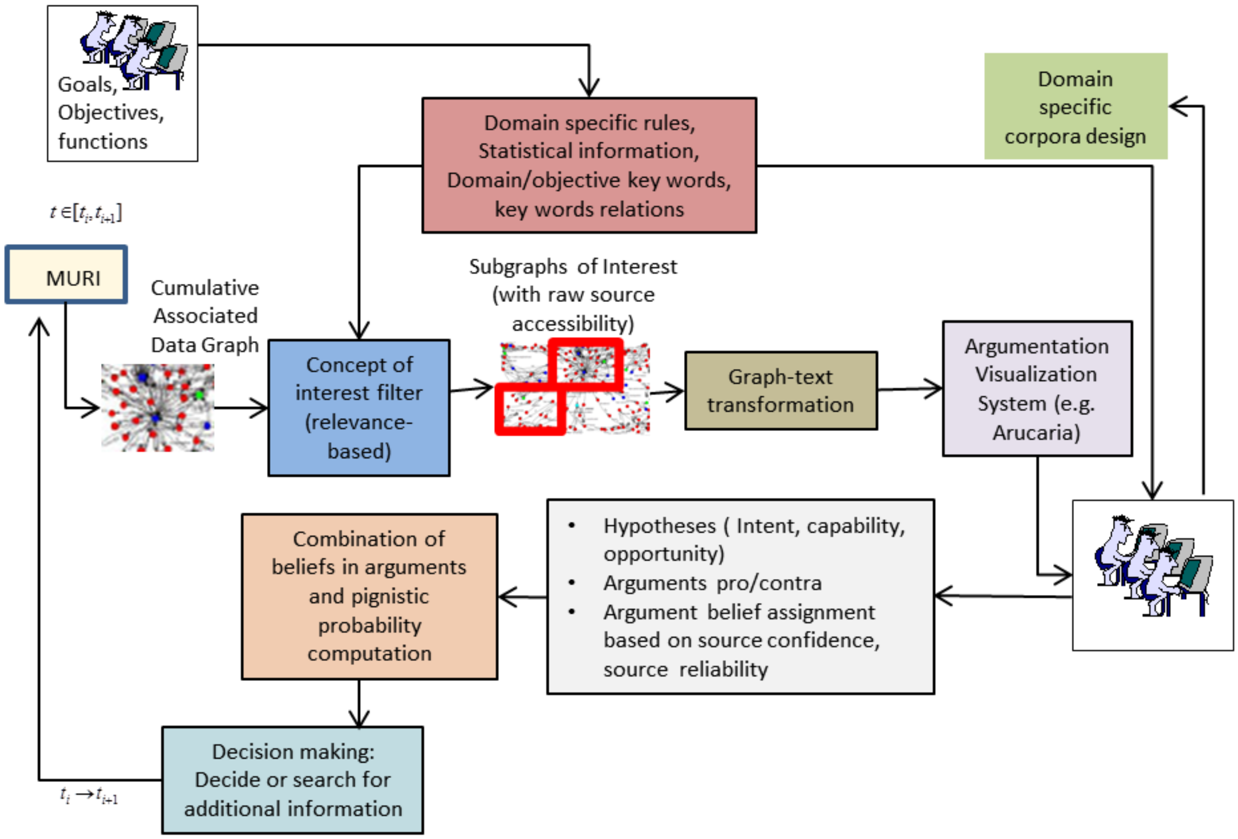


Figure 32: Building arguments, assigning beliefs (Human-based system)

In the human-based systems, relevant pieces of story in the cumulative associated sub graph are presented to analysts. An argumentation visualization system, e.g. Araucaria, offers users visualization of the argument structure of a text and assists them in the drafting of the argumentation structure of a text. In order to use the relevant associated data subgraph as input to such systems, a graph-text transformation process is required. Analysts then formulate arguments pro and contra and assign beliefs to them based on their experience, source confidence and source reliability represented by linguistic variables as well as domain specific statistics. The process of argument creation has to take into account not only a story supporting arguments pro one of the threat components but also a competing story supporting arguments contra threat hypotheses. These stories can come from domain specific rules, domain keywords, as well as domain keyword relations. Information produced by analysts serves as input in belief combination and decision making modules, which will be described in the next subsection.

One of the important subprocesses of both pure human and human-computer based systems is the process of domain corpora design. Domain corpora are repositories of structures such as sentences containing arguments pro, contra, and sentences not containing arguments, and dealing with threat hypotheses. Domain corpora are useful for training and testing argument mining classifiers in designing a more sophisticated human-computer argument creation system.

The human-computer system presented in Figure 33 below differs from the pure human system in the way the arguments are created. As opposite to the latter, it utilizes an automatic

sub-process of argument mining via either textual argument classification [G.2.31],[G.2.32] or via the graph matching methodology developed on the MURI program. The textual argument mining comprises of processes for argument detection and classification. The feasibility of textual argument mining has been shown in [G.2.31], in which argument detection and classification was conducted in the legal domain. Features used in [G.2.31] for argument classification and terminal and non-terminal symbols from the context-free grammar used in the argumentation structure detection are presented in Tables 1 and 2, respectively. In the human-computer systems the results of argument mining are presented to a human user for verification and belief assignment.

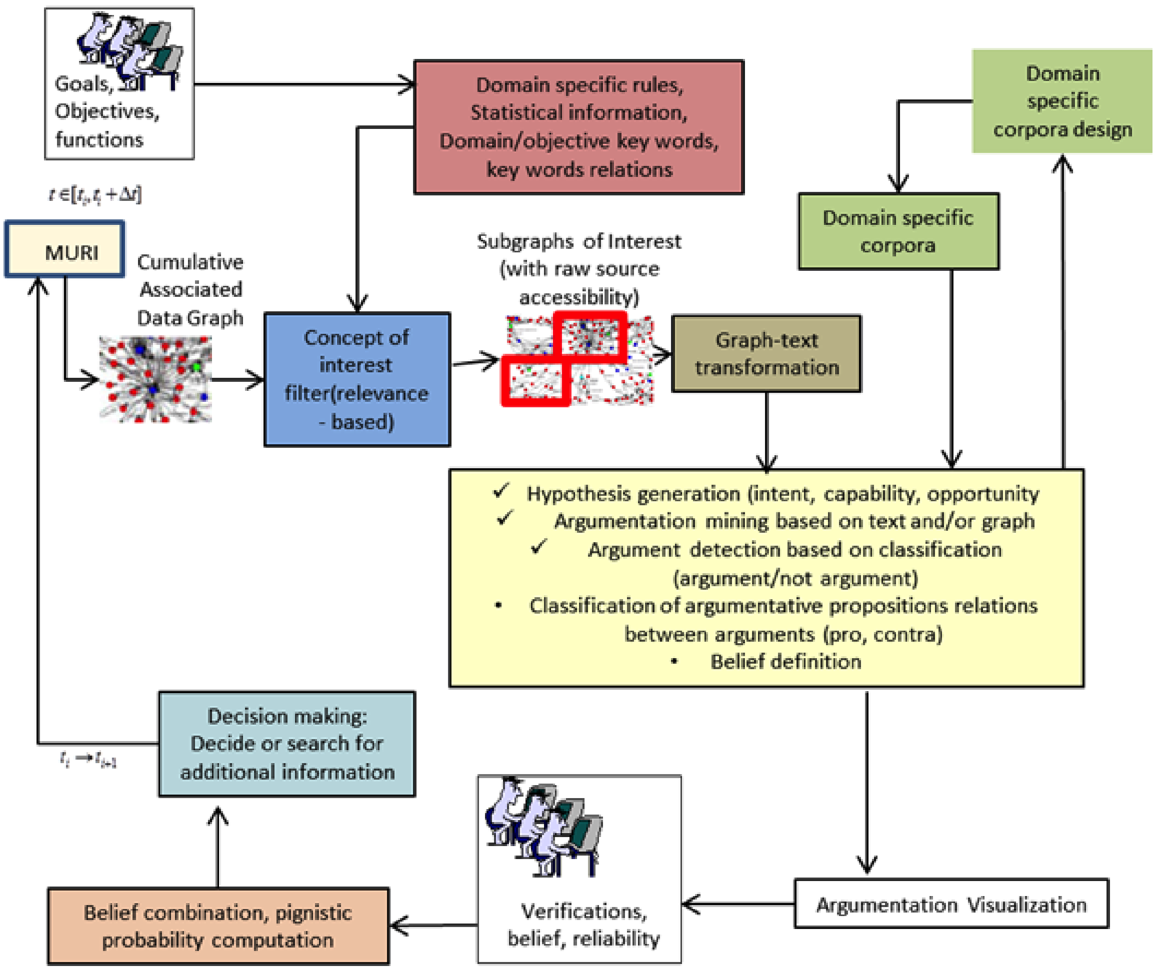


Figure 33: Building arguments, assigning beliefs (Human-computer system)

The third possible processing of argument construction is based on case (story)-based reasoning (see, e.g. [G.2.33],[G.2.34]). A functional diagram of story-based semiautomatic construction of arguments is presented in Figure 34. In case-based reasoning in general, knowledge is stored in a library of past cases, rather than in a knowledge base containing rules. In the story-based argumentation approach, knowledge is stored in a library of historical arguments supporting and refuting hypotheses about threat parts (capability, opportunity, and intent). It can be also augmented by domain related stories along with their descriptors as well as corresponding arguments.

This library is used for argument mining and retrieval of the best match for the concept of interest obtained by filtering the cumulative associated data graph. The process of argument retrieval requires designing a similarity measure between pieces of information contained in the cumulative associated data graph and arguments in the library. We propose this operation be implemented via the inexact stochastic graph matching methodology developed on this program, with template graphs derived from the historical library matched against the filtered cumulative associated data graph.

The similarity is measured on both feature and semantic levels, and expressed as belief computed as a function of the matching scores (i.e., the stochastic graph matching similarity score). The matching process produces a set of messages/parts of messages matching input pieces of stories as well as a set of possible arguments pro and contra. The retrieved similar stories and arguments are sent to an analysts' screen for adaptation, interpretations, and possible new stories construction, which are then send to the library of stories.

Table 23: Example of features used in [G.2.31] for argument classification

Absolute Location	Position of sentence absolutely in document; 7 segments
Sentence Length	A binary feature, which indicates that the sentence is longer than a threshold number of words (currently 12 words).
Tense of Main Verb	Tense of the verb from the main clause of the sentence; having as nominal values "Present", "Past" or "NoVerb".
History	The most probable argumentative category (among the 5 categories) of previous and next sentences).
Information 1st Classifier	The sentence has been classified as argumentative or non-argumentative by a first classifier.
Rhetorical Patterns	Type of rhetorical pattern ocurring on current, previous and next sentences (e.g. "however,"); we distinguish 5 types (Support, Against, Conclusion, Other or None).
Article Reference	A binary feature indicating whether the sentence contains a reference to an article of the law, detected with a POS tagger .
Article	A binary feature indicating that the sentence includes the definition of an article detected again with the help of a POS tagger .
Argumentative Patterns	Type of argumentative pattern ocurring in sentence; we have distinguished 5 types of patterns in accordance with our 5 categories (e.g. "see, mutatis mutandis,", "having reached this conclusion", "by a majority").
Type of Subject	The agent of the sentence is the applicant, the defendant, the court or other. The type of agent is detected with the POS tagger.
Type of Main Verb	Argumentative type of the main verb of the sentence; we distinguish 4 types (premise, conclusion, final decision or none), implemented as a list of corresponding verbs, which are detected in the text also with a POS tagger .

Table 24: Terminal and non-terminal symbols from the context-free grammar used in [G.2.31] in the argumentation structure detection

T	General argumentative structure of legal case.
A	Argumentative structure that leads to a final decision of the factfinder $A = \{a_1, ..., a_n\}$, each a_i is an argument from the argumentative structure.
D	The final decision of the factfinder $D = \{d_1, ..., d_n\}$, each d_i is a sentence of the final decision.
P	One or more premises $P = \{p_1, ..., p_n\}$, each p_i is a sentence classified as premise.
C	Sentence with a conclusive meaning.
n	Sentence, clause or word that indicates one or more premises will follow.
s	Sentence, clause or word neither classified as a conclusion nor as a premise ($s! = \{C\|P\}$).
r_c	Conclusive rhetorical marker (e.g. therefore, thus, ...).
r_s	Support rhetorical marker (e.g. moreover, furthermore, also, ...).
r_a	Contrast rhetorical marker (e.g. however, although, ...).
r_{art}	Article reference (e.g. terms of article, art. para. ...).
v_p	Verb related to a premise (e.g. note, recall, state,...).
v_c	Verb related to a conclusion (e.g. reject, dismiss, declare, ...).
f	The entity providing the argumentation (e.g. court, jury, commission, ...).

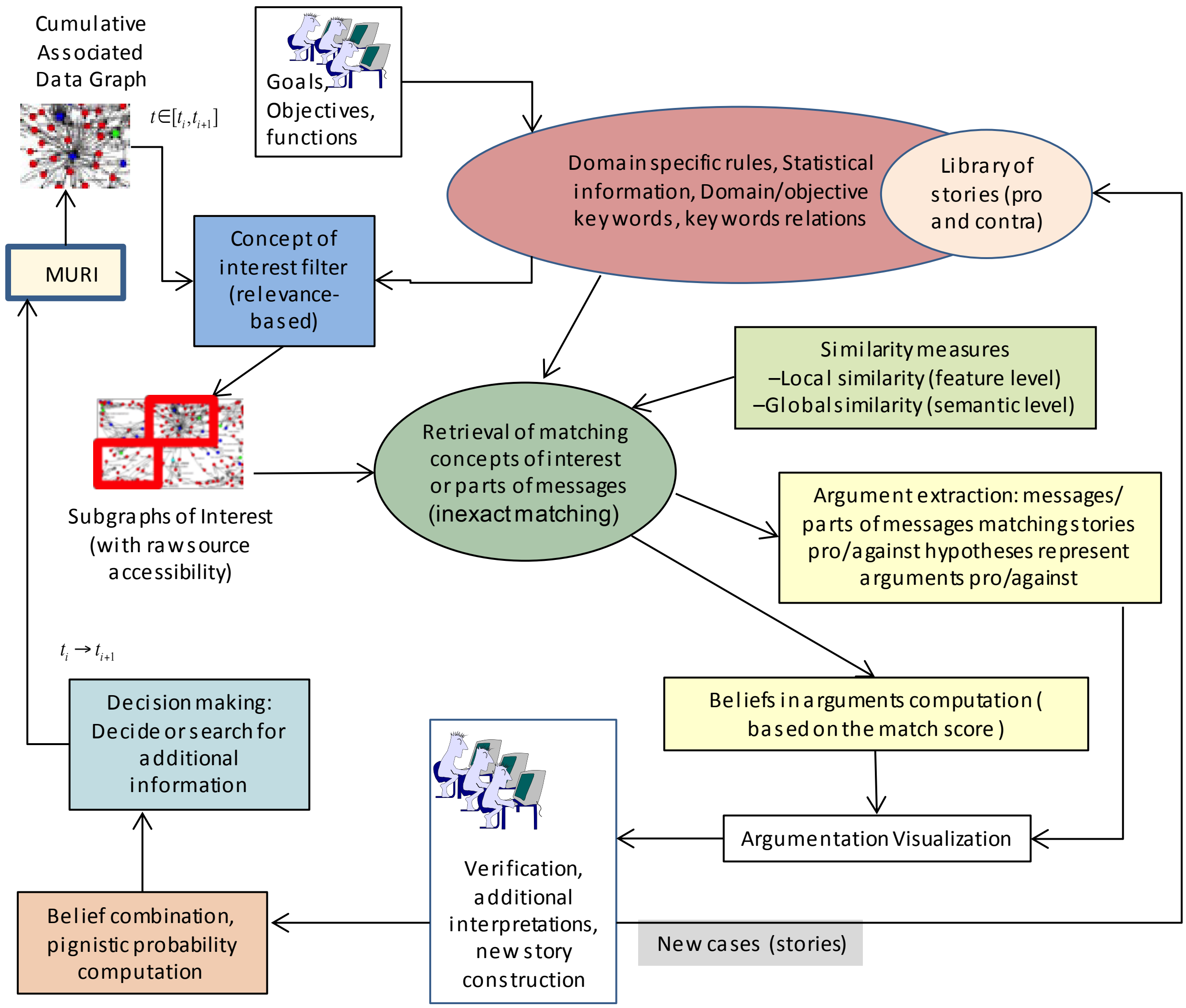


Figure 34: Story-based argument retrieval

3.1.7.2.3.2 Belief combination and decision making

Figure 35 presents a functional diagram for two sub-processes: belief combination and anytime decision making. Beliefs in each part of the threat (intent, opportunity and capability) are computed by combining beliefs in *pro* and *contra* arguments constructed during an episode *i* ($t \in [t_i, t_i + \Delta t]$) by the unnormalized Dempster rule of combination (Eq. 1). If during episode *i* arguments for all three parts were present, the unnormalized Dempster rule of combination is used to fuse these combined beliefs to obtain belief into imminent threat, which is then transformed into *pignistic probability* (Eq. 4) for decision making. This *pignistic probability* is then compared with a predefined domain-specific time-varying threshold to decide whether to select the "imminent threat" hypothesis [G.2.5],[G.2.6] as valid. The threshold is a decreasing function of time and shaped to encourage early decisions, while incorporating a finite decision deadline. If the threshold is satisfied, an alert is sent to the decision maker to verify the presence of threat. Otherwise, the pairwise belief combination for each pair of threat parts is performed to obtain beliefs in potential threat, which are then transformed into *pignistic probability* (Eq. 4). Each

pairwise pignistic probability is compared with a domain specific potential threat threshold. If the threshold is satisfied an alert is sent to the decision maker, otherwise additional information is searched for during the next time interval, the length of which is domain specific (e.g., based on sampling frequency).

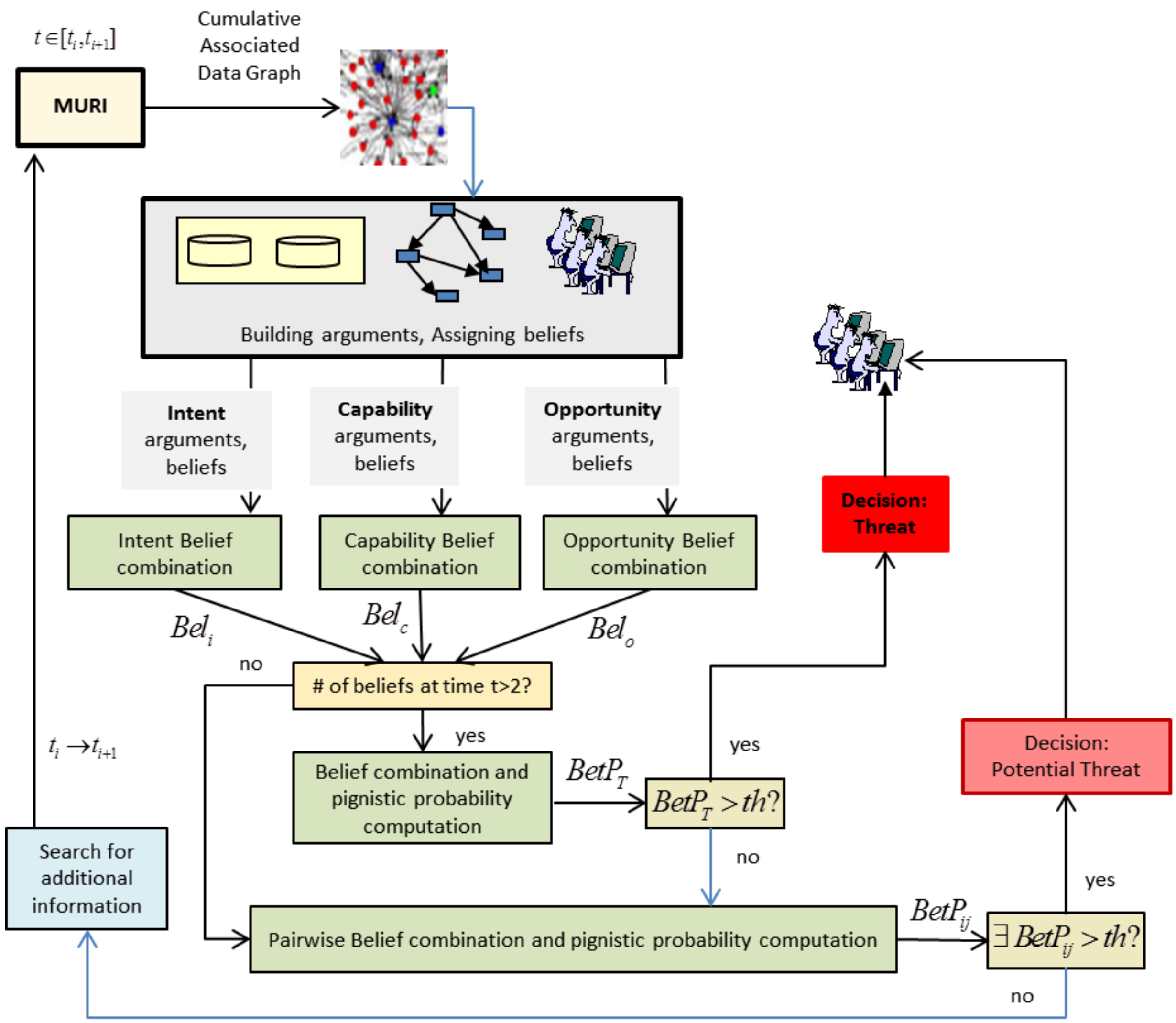


Figure 35: Belief computation and decision making

Similarly, if during episode *i* arguments for only two of these three parts were present, the pairwise belief combination is performed to obtain beliefs into potential threat, which are then transformed into *pignistic probability* (Eq. 4). Each pairwise pignistic probability is compared with a domain and potential threat specific threshold. If the threshold is satisfies an alert is sent to the decision maker, otherwise additional information is searched for. Additional information is also searched if arguments for only one threat part are present.

3.1.7.2.4 Conclusions

This progress report summarized the result of research conducted to perform a review of existing argumentation related publications and available software as well as to define a go-forward approach to Belief-based Argumentation for holistic situation/threat assessment, and a functional diagram of this approach. The Story-based Belief-based Argumentation process defined in this research combines the Transferable Belief Model, story-based argumentation, graph matching, and anytime decision making. The implementation of this approach faces many challenges that include but are not limited to designing:

- a library of domain specific stories, arguments, rules, and key words
- methods for *pro* and *contra* argument mining
- specific methodology of defining beliefs to and reliabilities of arguments as well as evaluation and incorporation into the processing argument importance
- graph-text transformation or extending argumentation visualization systems to present a simultaneous graph and text display
- investigating the integration of graph analytic output such as high value individual identification from social network analysis or link analysis results into the formation of arguments

Another challenge is to adopt the existing software developed for argumentation visualization to implement particular argumentation schemes, e.g.. Araucaria, to be used for belief-based argumentation.

3.1.7.2.5 References

[G.2.1] P. Smets, Data Fusion in the Transferable Belief Model, In: *Proc. of the FUSION'2000-Third Conference on Multisource- Multisensor Information Fusio*n, 2002, pp. 21-33.

[G.2.2] F. Delmotte, Ph.Smets, Target Identification Based on the Transferable Belief Model Interpretation of Dempster-Shafer Model., *IEEE Trans. Syst., Man, Cybern. A*: 34, 457-471, (2004)

[G.2.3] B. Ristic, Ph. Smets, Ph. (2004) Kalman filters for tracking and classification and the transferable belief model. In: *Proc. of the FUSION'2004-Seventh Conference on Multisource Information Fusion,* July 2004.

[G.2.4] J. García, J. Gómez-Romero, M.A. Patricio, J.M. Molina, G. Rogova, Representation and Exploitation of Context Knowledge in a Maritime Surveillance Scenario, in: Proc. of *FUSION'2010-14th Conference on Multisource Information Fusion*, Chicago, IL, 2011.

[G.2.5] G. Rogova, P. Scott, C. Lollett, R. Mudiyanur, Reasoning about situations in the early post-disaster response environment, in: *Proc. of the FUSION'2006-9th Conference on Multisource Information Fusion*, 2006.

[G.2.6] G. Rogova, Adaptive real-time threat assessment under uncertainty and conflict, 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Antonio, TX; 03/2014

[G.2.7] S. Zilberstein, "Using Anytime Algorithms in Intelligent Systems". AI Magazine Vol. 17, 3, 1966

[G.2.8] G. Rogova, P. Scott, C. Lollett, Distributed Reinforcement Learning for Sequential Decision Making, In: Proc. of the FUSION'2002-Fifth Conference on Multisource Information Fusion, 2002.

[G.2.9] G. Rogova, M. Hadrazagic, M-O. St-Hilaire, M. Florea, P. Valin, Context-Based Information Quality for Sequential Decision Making, in *Proc. of the 2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support* ,2013.

[G.2.10] G. Rogova, E. Bosse, Information quality in information fusion, in: Proc. of *FUSION'2010-13th Conference on Multisource Information Fusion*, Edinburgh, Scotland, 2010.

[G.2.11] G. Simari, I. Rahwan, *Argumentation in artificial intelligence,* Springer, 2009

[G.2.12] T.J.M. Bench-Capon, P.E. Dunne , Argumentation in Artificial intelligence*, Artificial Intelligence 171 619–641, 2007.*

[G.2.13] P. M. Dung, (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games, *Artificial Intelligence* 77 (2), pp. 321–357.

[G.2.14] Rahwan (Ed.) (2005). *Special Issue on Argumentation in Multi-Agent Systems, Journal of Autonomous Agents and Multiagent Systems*. Vol 11, No 2.

[G.2.15] F.J. Bex (2013) Abductive Argumentation with Stories. *ICAIL-2013 Workshop on Formal Aspects of Evidential Inference.*

[G.2.16] F. BEX ,H PRAKKEN Investigating Stories in a Formal Dialogue Game, In *Proceedings of COMMA 2008*. Amsterdam etc, IOS Press 2008

[G.2.17] H, Prakken, Logical Tools for Modelling Legal Argument. *A Study of Defeasible Argumentation in Law and Philosophy Library*, Dordrecht/Boston/London: Kluwer Academic Publishers (1997).

[G.2.18] N.M. Mirza, A-N Perret-Clermon, Argumentation and Education: Theoretical Foundations and Practices, *Springer Science & Business Media*, Jun 19, 2009.

[G.2.19] J. Lowrance, I. Harrison, A. Rodriguez, E. Yeh, T. Boyce, J. Murdock, J. Thomere, and K. Murray, Template-based Structured Argumentation, in: *Knowledge cartography*, Springer, 2008.

[G.2.20] Bondarenko, A., Dung, P., Kowalski, R., and Toni, F, An abstract, argumentation-theoretic approach to default reasoning, *Artificial Intelligence*, 93, 63–101,1997.

[G.2.21] P. M. Dung, R. Kowalslski, F. Toni, Assumption-based argumentation, in: G. Simari, I. Rahwan *Argumentation in Artificial Intelligence,* Springer, 2009, pp 199-218 .

[G.2.22] G. Vreeswijk, "Abstract argumentation systems," *Artificial Intelligence*, 90, 225–279, 1997.

[G.2.23] J. Pollock, J. (1987), Defeasible reasoning, *Cognitive Science*, 11, 481–518, 1987.

[G.2.24] J. Pollock, Justification and Defeat, *Artificial Intelligence*, 67, 377–408, 1994

[G.2.25] R. Haenni, J. Kohlas, N. Lehmann. Probabilistic Argumentation Systems, in: J. Kohlas, S. Moral (eds). *Handbook of Defeasible Reasoning and Uncertainty Management Systems,* Vol. 5, Kluwer, 2001.

[G.2.26] B. Verheij, Artificial argument assistants for defeasible argumentation, Artificial Intelligence 150, 291–324, 203.

[G.2.27] http://www.computing.dundee.ac.uk/staff/creed/araucaria/

[G.2.28] Bex, F.J. *Arguments, Stories and Criminal Evidence:A Formal Hybrid Theory*, Springer, Dordrecht.

[G.2.29] F.J. Bex , Abductive Argumentation with Stories. *ICAIL-2013 Workshop on Formal Aspects of Evidential Inference, 2013.*

[G.2.30] E. Little, G. Rogova, An Ontological Analysis of Threat and Vulnerability, in: *Proc. of the FUSION'2006-9th Conference on Multisource Information Fusion*, 2006.

[G.2.31] Palau, Moens, Argumentation Mining: The Detection, Classification and Structure of Arguments in Text, *ICAIL-2009* Barcelona, Spain.

[G.2.32] Wyner, A., Mochales-Palau, R., Moens, M. F., & Milward, *Approaches to text mining arguments from legal cases* (pp. 60-79). Springer Berlin Heidelberg, 2010.

[G.2.33] A. Aamodt, E. Plaza, Case-based reasoning: foundational issues, methodological variations and system approaches, AI Communications 7 (1) 39–59, 1994.

[G.2.34] S. Heras, J. Jordán, V.Botti, V. Julián, Argue to agree: A case-based argumentation approach, International Journal of Approximate Reasoning 54 (2013) 82–10

3.2 University of Illinois at Urbana-Champaign (UIUC)

3.2.1 Hard+Soft Data Association

Data gathered during various Counterinsurgency (or COIN) operations is in different formats. For example, the data gathered by human observers in the form of field reports, notes, journals etc., contains structured or unstructured natural language, which is also called soft data. On the other hand, the data gathered by human operated or automated physical sensors, such as cameras, LIDAR, acoustic sensors, is called hard data. These data usually contain references to entities and their relationships, describing various attributes of each. These data form the basis for sense-making and situational awareness purposes, enabling a better understanding of the state of the real world. Many times, multiple references in the observed data represent the same real world entity. This duplication might stem from additions to the data over time, typographical errors, or multiple data entries. These duplicate references potentially limit the efficiency of the database and might cause problems like incorrect information retrieval and wasted storage space. The role of data association is to identify and merge the references which correspond to the same real world entity, forming fused (cumulative) evidence. This cumulative evidence will contain more information about the real world entities than offered by any single observation and it can be used in sense-making tasks, to build hypotheses or draw conclusions on the current state of the real world.

The cumulative evidence obtained from the data association task needs to be evaluated in fair and objective manner, in order to make sure that it correctly reflects the state of the real world. Additionally, this evaluation of the cumulative evidence needs to be accomplished with minimal human intervention, so as to make this task more efficient. Therefore during the Year 5 of the MURI program, the efforts were primarily focused on development of an automated testing and evaluation methodology for gauging the performance of data association [H.1].

3.2.1.1 Scientific / Technical Accomplishments

Years 1, 2, and 3

- Development of Graph association model for the association of richly relational data.
- Implementation of the Distributed ("Cloud") version of the Lagrangian heuristic.
- Development of the Incremental association approach for streaming datasets.

Year 4

- Development of ground truth for 114 soft messages of the Sunni Criminal Thread (SUN).
- Incorporation of Incremental association approach into MURI architecture.
- Deployment of Incremental association algorithm in a networked scenario demonstration.

Year 5

- Development of ground truth for 13 hard messages of the Sunni Criminal Thread (SUN).
- Development automated testing and evaluation methodology for data association.
- Comparison of various data association formulations and algorithms in objective manner, in terms of accuracy and execution time.

3.2.1.2 Overview of Data Association Process

The data association process is divided into different steps as seen in Figure 36. Initially, the hard and soft messages are converted into relational, attributed graphs which are used as an input to the data association engine. The soft messages are processed using the NLP tool Tractor [H.2], while the hard messages are processed using machine learning based detection and tracking algorithms.

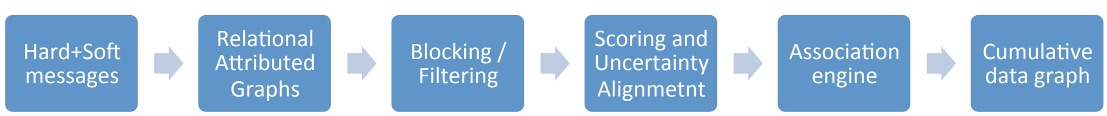


Figure 36: Association process overview

Next, a pairwise comparison is conducted between the nodes and edges in the graphs and a similarity score is calculated for each pair using various string similarity measures. In a graph with modest size there are $O(n^2)$ pairs of nodes (or edges), and scoring each pair is asymptotically prohibitive even when utilizing a relatively efficient scoring function. Therefore, blocking or filtering techniques are used to avoid scoring the element pairs which are likely to have very low scores. Blocking is the process of selecting the candidates for scoring, based on some higher level criteria. If a particular node or edge pair meets these criteria, they are scored

using the scoring function; otherwise they are not included in the association problem, thereby reducing the size of the association instances. In this application, the "type" attribute was used as the blocking criteria for selecting the node pairs. For example, a node of the type "person" was scored only with another node of the type "person," and not with a node of the type "location." This way many of the node pairs were eliminated, significantly reducing the runtime of the scoring process. Currently in our application, edges do not possess such higher level criteria for blocking and therefore, all the edge pairs were considered for scoring.

Once particular node/edge pairs pass the blocking criteria, their similarity score is calculated by comparing the values of their various attributes. Depending on the type, each node (or edge) may have a specific set of attributes. For example, the attribute set of a "person" node typically contains type, name, sex, religion, race, age, height, and weight. This set is sometimes referred to as the feature vector of the node (or edge). Some of the features from the feature vector have textual values (e.g. type, name, race, etc.), while others have numerical values (e.g. age, height, weight, etc.). The values of the like features are compared using a feature specific scoring function and a feature level similarity score is generated. For scoring the textual features, various string similarity functions were used, such as Levenshtein distance [H.3] and semantic similarity calculations within the WorldNet Similarity Library [H.4]. Numerical features were scored using direct comparison. The total similarity score between a node/edge pair was calculated as the weighted sum of individual feature scores. The feature weights were obtained by training a logistic regression model on the feature scores. The nodes and edges also undergo the process of uncertainty alignment [H.5] [H.6]; which captures and quantifies the ambiguity of the attribute values. The ambiguous attribute values are modeled using probabilistic or possibilistic distribution functions and the similarity score for such attributes is calculated as a function of the probability values obtained from the respective distributions.

After scoring the node and edge pairs, the data association problem is formulated and solved using one of the algorithms described in the next section. The output of the data association engine is a cumulative data graph, in which the pairs of associated entities are merged together, along with their relationships. This cumulative data graph can be considered as fused evidence and it can be used in the various downstream analyses like graph matching or social network extraction.

3.2.1.3 Formulations and Algorithms

Given a set of relational attributed graphs; and the similarity scores between pairs of nodes and pairs of edges, data association tries to maximize the total similarity score by clustering (or associating) the similar nodes/edges across the input graphs. The attributes of the associated nodes (or edges) are fused (merged), to produce a "cumulative data graph" (CDG), representing cumulative situational evidence. Data association problem can be modeled as a graph association problem. Depending on the number of graphs, several mathematical formulations can be obtained, and they can be morphed into one another by addition or deletion of various constraints. Figure 37 depicts the relationships between these different formulations. We have studied three of these formulations: GA^N, MDADC, and CPP, which are described in detail in the subsequent sections.

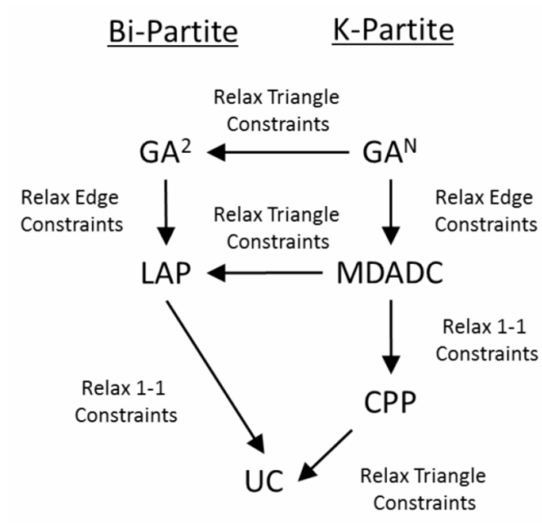


Figure 37: Taxonomy of data association formulations

3.2.1.3.1 Graph Association (GA^N)

The data association problem can be modeled as a graph association problem (GA^N) [H.7], which is a generalization of the multi-dimensional assignment problem. An integer programming (IP) formulation was developed for the GA^N [H.8]. This formulation has four types of constraints (in the presented order): node association, edge association, node transitivity, and edge transitivity. The complete formulation can be written as follows.

$$
\begin{aligned}
GA^N = \max \sum_{v_i \in V} \sum_{v_j \in V} A_{ij} x_{ij} + \sum_{e_{ik} \in E} \sum_{e_{jl} \in E} B_{ikjl} y_{ikjl} & \\
\text{subject to: } \sum_{v_i \in V_{G_k}} x_{ij} \le 1 \qquad & \forall G_k \in G, G_{k'} \in G \backslash G_k, v_j \in G_{k'} \\
2y_{ikjl} \le x_{ij} + x_{kl} + x_{il} + x_{kj} \qquad & \forall y_{ikjl} \in E \times E \\
x_{ij} + x_{ik} \le x_{jk} + 1 \qquad & \forall (v_i, v_j, v_k) \\
y_{ijkl} + y_{ijmn} \le y_{lkmn} + 1 \qquad & \forall (e_{ij}, e_{kl}, e_{mn}) \\
y_{ikjl}, x_{ij} \in \{0, 1\} &
\end{aligned}
$$

GA^N is also a generalization of the quadratic assignment problem (QAP), which is an NP-hard problem. Therefore, no existing polynomial time algorithm can solve this problem optimally within a guaranteed time limit, and we have to rely on efficient heuristics, which can provide good enough solutions in permissible time. For this purpose, a Lagrangian heuristic was developed, in which the node and edge transitivity constraints are introduced into the objective function using appropriate dual multipliers. The dual multipliers are adjusted in each iteration, so as to minimize the penalty incurred by infeasible constraints. Thus the algorithm obtains better and better solutions until a provably optimal solution is found, a pre-determined optimality gap is achieved, or the permissible time limit has been exceeded. This heuristic is able to solve small and medium sized data association problems within 3% of the optimality.

3.2.1.3.2 Multidimensional Assignment Problem with Decomposable Costs (MDADC)

The IP formulation of the GA^N contains a large number of constraints and variables, even for small or medium sized graphs. If the number of graphs or the numbers of nodes/edges in each

graph are sufficiently large, then the above mentioned procedure can become extremely time-consuming. Therefore a new method needed to be developed which could potentially divide the work into multiple processors and alleviate the computational burden. The GA^N formulation is quite complex in the sense that it is not conducive to parallelization. To address this problem, a relaxed version of the data association formulation was developed [H.9]. This formulation is called multi-dimensional assignment problem with decomposable costs (MDADC) [H.9], and it is obtained by removing the edge association and edge transitivity constraints from GA^N. The complete formulation can be written as follows.

$$
\begin{aligned}
\text{MDADC} = \max & \sum_{v_i \in V} \sum_{v_j \in V} A_{ij} x_{ij} \\
\text{subject to :} & \sum_{v_i \in V_{G_k}} x_{ij} \leq 1 \qquad \forall G_k \in G, G_{k'} \in G \backslash G_k, v_j \in G_{k'} \\
& x_{ij} + x_{ik} \leq x_{jk} + 1 \qquad \forall (v_i, v_j, v_k) \\
& x_{ij} \in \{0, 1\}
\end{aligned}
$$

This problem is easier to parallelize than GA^N, due to the absence of the complicating edge constraints. A new parallel version of the Lagrangian heuristic was developed and implemented using the Map/Reduce programming architecture. In this algorithm, the node transitivity constraints are penalized with the help of dual multipliers. The resulting problem can be decomposed into multiple linear assignment sub-problems (LAPs), which can be solved in distributed fashion by multiple processors. This algorithm shows significantly faster execution times and good scalability behavior for problems containing up to 30,000 nodes.

3.2.1.3.3 Clique Partitioning Problem (CPP)

The two formulations described above rely on the fact that the input data is static and all the data points are available at the runtime. They also assume that the within message co-referencing is perfect, i.e. each unique entity has at most one mention in any given message. However, these assumptions might not hold for the data obtained from the real world, which could be frequently changing. These changes might correspond to addition of new entities into the dataset or modifications in the attributes of the existing entities. This dynamism could potentially pose a question to the verity of the cumulative evidence that was obtained in the previous state of the system, unless the changes are propagated into the cumulative evidence. One approach to deal with this problem is to re-execute one of the above data association algorithms on the entire dataset for the current state. However, this approach can be inefficient over a long enough time horizon, as the dataset can become extremely large. For this purpose, this problem was modeled as a clique partitioning problem (CPP) [H.11] [H.12] [H.12], by further removing the node association constraints from the MDADC formulation. The CPP formulation only contains the node transitivity constraints. The complete formulation can be written as follows.

$$
\begin{aligned}
\text{CPP} = \max & \sum_{v_i \in V} \sum_{v_j \in V} A_{ij} x_{ij} \\
\text{subject to :} \ & x_{ij} + x_{ik} \leq x_{jk} + 1 \qquad \forall (v_i, v_j, v_k) \\
& x_{ij} \in \{0, 1\}
\end{aligned}
$$

Based on this formulation, a new sequential algorithm was developed which can handle new additions to the data as well as incremental changes to the existing data over a period of time. This algorithm considers each node of the newly arrived graph and scores it against the clusters from the previous data association solution. There are three possibilities: (1) the newly arriving

node is not associated with any existing nodes, in which case it forms its own cluster; (2) the newly arriving node is added to an existing cluster, and remaining clusters are unaltered; and (3) the newly arriving node is added to an existing cluster, with possible restructuring of some of the other clusters. The algorithm was tested on synthetic as well as real world datasets and it was shown to provide much better results as compared to other competing algorithms/heuristics. The main feature of this algorithm is that it can potentially recover the incorrect associations that were made in the past, which makes it suitable for noisy input data.

3.2.1.4 Evaluation of Data Association

The evaluation methodology for data association is divided into two main parts: ground truth development and evaluation algorithm, as described below.

3.2.1.5 Ground Truth Development

Development of the ground truth is a key step for evaluating the performance of any data association algorithm. The ground truth is typically prepared by one or more human analysts and it represents the answer key to the data association solution, against which the association algorithm is graded. Individual ground truths are developed for the hard and soft messages as described below.

3.2.1.5.1 Soft Data Ground Truth

The soft ground truth is made up of two lists: (1) list of "unique entities" and (2) list of "observed entities." In the "unique entities" list, the entities appearing in all of the soft messages are listed and a unique entity identifier (UID) is assigned to each one of them. In "observed entities" list, all the mentions of a particular unique entity are listed under the specific UID, along with the message number in which they were observed. For preparing the soft ground truth, each text message is carefully read and understood by the analyst, and all the entities are identified along with their types (e.g. person, location, organization, etc.). For each of the identified entities, it is determined whether that entity is being encountered for the first time or it has been encountered before, in some previous text message. If the entity is encountered for the first time it is added to the "unique" list with a new UID; and since it also counts as a mention, it is added to the "observed" list, under the same UID. Any subsequent mentions of that entity are added to the "observed" list, under the UID of that entity. In addition to the entity names, the analyst also records the pedigree information of the observed entity, which contains the starting character position and the total number of characters in the textual description of that particular observation. The pedigree information serves an important role in automating the evaluation process. An example of the soft ground truth is shown in Figure 38. Within the SUN message set there are 140 multiply-mentioned entities, with a total of 1,024 mentions across the 114 soft messages.

Unique ID	Entity Name	Entity Type
1	Iraqi Domestic Counter Intelligence	Organization
2	Sunni	Organization
3	Rashid	Location
4	Coalition Forces	Organization
⋮	⋮	⋮
7	Rashid Criminal Group	Organization
⋮	⋮	⋮
9	Dhanun Ahmad Mahmud	Person
⋮	⋮	⋮

Entity ID	Message ID	Entity Name	Token Start Position	Token Length
1	54	Iraqi Domestic Counter Intelligence	15	35
2	54	Sunni	85	5
3	54	Rashid	121	6
4	54	Coalition Forces	131	16
3	54	Rashid	207	6
⋮	⋮	⋮	⋮	⋮
2	59	Sunnis	75	6
9	59	Dhanun Ahmad Mahmud	121	19
⋮	⋮	⋮	⋮	⋮
9	64	his	157	3
⋮	⋮	⋮	⋮	⋮
7	143	Sunni Criminal Group	241	20
⋮	⋮	⋮	⋮	⋮

Figure 38: Association Ground Truth: (a) unique entities (b) observed entities

3.2.1.5.2 Hard Data Ground Truth

The hard data ground truth consists of the entities (e.g. persons and vehicles) within the visual data source (video) and information regarding the bounding boxes of those entities, within different frames of the video, as seen in Figure 39. This bounding box information is preserved across each frame of the video, creating a ground truth bounding box track. The research on incorporating this type of ground truth into the data association evaluation framework is still ongoing, and we have not used this type of ground truth in evaluating data association performance.

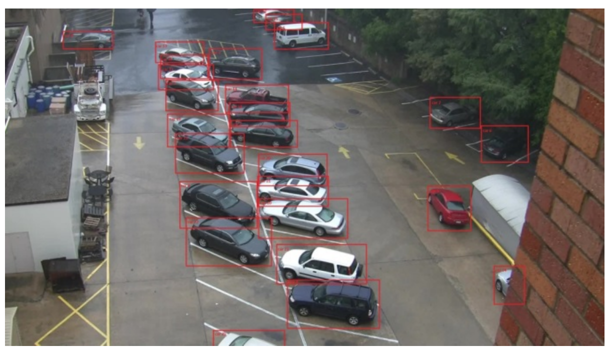


Figure 39: Hard data bounding box ground truth example

We have created a simpler version of the ground truth for the hard messages, in which entities are considered at the message level, rather than frame or bounding box level. For each video from the hard message set, the unique and observed entities are identified and listed in the similar fashion as the soft ground truth. Additionally, the hard messages are compared to each of the soft messages, for detecting cross-modality commonalities. If the analyst deduces that a hard message corresponds to a particular soft message, then the individual entities are compared and appropriate names and UIDs from the soft ground truth are assigned to them. Analogous to the soft ground truth, the analyst creates dummy pedigree information for each observation so as to make it compatible for automated evaluation. Within the SUN message set there are 30 multiply-mentioned entities, with a total of 42 mentions across the 13 hard messages.

3.2.1.6 Evaluation Metrics

By executing one of the data association algorithms on the hard and soft input data, a cumulative data graph (CDG) is obtained. This CDG is programmatically compared with the ground truth and three types of entity pairs are counted: (a) correctly associated, which contains

node pairs that are merged in the CDG and have the same UID in the ground truth; (b) incorrectly associated, which contains node pairs that are merged in the CDG but do not have the same UID in the ground truth; and (c) incorrectly not associated, which contains node pairs that are not merged in the CDG but have the same UID in the ground truth. These counts are used to calculate the following three metrics for quantifying the performance of data association:

- Precision: This is the ratio of correctly associated entity pairs to the total number of associated entity pairs, i.e. $\left(\frac{a}{a+b}\right)$. This value lies between 0 and 1.
- Recall: This is the ratio of correctly associated entity pairs to the total number of correctly associated and incorrectly not associated entity pairs, i.e. $\left(\frac{a}{a+c}\right)$. This value also lies between 0 and 1.
- F-score: This is the harmonic mean of the Precision and Recall values, i.e. $\left(\frac{2PR}{P+R}\right)$.

Together, the Precision, Recall, and F-score represent the accuracy of the association results; with higher values typically indicating greater accuracy. Precision and Recall are often competing with each other, i.e. if we configure the algorithm to maximize the Precision, then it might produce poor value for the Recall. For this reason the algorithms in general are configured to maximize the F-score, which tries to strike a balance between the Precision and Recall.

3.2.1.7 Automated Evaluation

3.2.1.7.1 Pedigree Records

The pedigree record for an entity/relationship is a tuple of three integers, which is used for recording the exact location of that entity in a particular text message. A pedigree record is composed of the message number containing that particular entity mention, and the starting character position and the number of characters in its textual description. For example, the entity "Dhanun Ahmad Mahmud" is mentioned in message 59, starting at character position 121, and it is 19 characters long. So the pedigree record tuple for this particular entity will be <<<59, 121, 19>>>. The pedigree records for the entities in the text messages are automatically identified during the natural language processing task and retained throughout data association and other downstream processes. During the development of the ground truth, the analyst records the pedigree information for all the entity mentions. During the data association process, the pedigree records of the associated entities are merged along with other attributes. Therefore, the pedigree information of the merged nodes can be compared programmatically with the pedigree information recorded in the ground truth to obtain the pertinent counts of correctly and incorrectly merged entities, which can be used to calculate the Precision, Recall and F-score.

3.2.1.7.2 Evaluation Algorithm

1. Initially, all the pedigree records are extracted from the CDG and also from the ground truth document, into two separate lists.
2. From the CDG pedigree record list, those instances are removed in which the character string is subsumed by a longer character string of some other pedigree record, present within a common graph element of the same message. For example, consider the two mentions "Dhanun Ahmad Mahmud" <<<59, 121, 19>>>, and "Ahmad" <<<59, 128,

5>>>. Both mentions are from the same message 59; represent the same entity; and the latter is subsumed by the former. In this case, pedigree record <<<59, 128, 5>>> is removed from the list. This step is necessary to avoid multiple counts of the same entity and potential inflation of the Recall value.

3. Next, the CDG pedigree records are compared with those from the ground truth and they are retained only if there is a partial or complete overlap with one of the pedigree records in the ground truth; otherwise they are removed. The surviving pedigree records represent the CDG entities that have a corresponding entity in the ground truth, and only those entities are considered for evaluation.
4. Next, two types of identifiers: EID and UID, are determined for each of the surviving pedigree records. EID is the ID of the CDG node which contains that pedigree record. UID is same as that of the corresponding ground truth entity, which can be obtained during the execution of Step 3.
5. For each pair of pedigree records, the respective values of EID and UID are compared with each other to determine whether they are correctly associated or not, and the appropriate counts "a," "b," or "c," (from Section 3.2.1.6) are incremented. Assuming that (EID1, UID1) and (EID2, UID2) are the identifiers for a pair of pedigree records (PR1, PR2), then Table 25 lists the various conditions and the associated inferences about the correct and incorrect association of the corresponding entities.
6. Next, all the pedigree records extracted from the ground truth which do not have a corresponding pedigree record in the CDG and the corresponding unique entities/mentions are identified. These represent the entities that should have been associated but they are not, because of the missing pedigree information in the propositional graphs. The pairs of mentions for each of those entities are counted and then added to the "incorrectly not associated" count (or "c" from Section 3.2.1.6).
7. Finally, the Precision, Recall and F-score are calculated using the counts obtained above.

Table 25: Identifier values and inferences

Identifier Values		**Inference**
$EID_1 = EID_2$	$UID_1 = UID_2$	Correctly associated (a := a + 1)
$EID_1 = EID_2$	$UID_1 \neq UID_2$	Incorrectly associated (b := b + 1)
$EID_1 \neq EID_2$	$UID_1 = UID_2$	Incorrectly not associated (c := c + 1)
$EID_1 \neq EID_2$	$UID_1 \neq UID_2$	Correctly not associated

There are several imprecisions stemming from the natural language processing step, such as incorrect entity types, and incorrect within message co-referencing. These could cause some

imprecision in counting the correctly and incorrectly associated entity pairs. Some of these issues can be tackled using the type restricted evaluation, as explained in the next section; while others are not so easy to deal with. Modeling data association as clique partitioning problem (CPP) and solving it using the Streaming Entity Resolution algorithm can also help recover some of the errors caused due to missing within message co-references, which potentially improves the F-score.

3.2.1.8 Type Restricted Evaluation

During the natural language processing step performed by Tractor, some of the entities could be assigned an incorrect type, as compared to the true type identified in the ground truth document. This imprecision may result in improper data association gating, reflected by missed or incorrect associations in the data association results. For example, assume that one of the messages contains a mention of the city named “Rashid,” which has the type “Location.” In some other message, there is a “Person” entity with the same name “Rashid.” If, during NLP step, the type of the former entity is incorrectly identified as “Person,” then data association might merge the two entities into a single entity; resulting in a wrong conclusion, and decreased Precision and F-score (a similar example can be provided for Recall). Therefore, to calculate the true Precision and Recall of data association, the entities which are incorrectly associated/not associated due to the type identification errors in the NLP, need to be discounted. This can be accomplished using type restricted evaluation, as described below:

1. Initially, for the CDG pedigree record pair under comparison, soft message entities, ground truth entities, and their respective types are identified.
2. If the two entities are incorrectly associated (or incorrectly not associated), then count “b” (respectively count “c”) are incremented, if and only if, at least one of the following two conditions holds:
    a. The “types” of both the entities are properly identified and they match with the “types” of the corresponding ground truth entities.
    b. The “type” of only one of the entities is properly identified and it matches with the “type” of the corresponding ground truth entity; while the “type” of the other entity is missing, both in the message as well as in the ground truth.
3. For the evaluation to be fair, the correct associations which overcame the incorrect type identification are disregarded. For two entities which are correctly associated, count “a” (from Section 3.2.1.6) is incremented, if and only if at least one of the conditions (i) or (ii) stated above hold. This prevents the unfair inflation of the Precision, Recall, and F-score.

In this way, the effect of incorrect type identification can be nullified and the Precision, Recall, and F-score of the data association solution can be potentially improved, using type restricted evaluation.

3.2.1.9 Testing

The evaluation strategy was tested on the three data association formulations and corresponding algorithms: sequential Lagrangian heuristic for GA^N, Map/Reduce Lagrangian heuristic for MDADC and streaming entity resolution algorithm for CPP, and their accuracy and

computing time were compared. The data association algorithms and the evaluation procedure were coded in Java and executed on Intel Core 2 Duo processor, with 3 GHz clock speed and 4 GB RAM. The statistics related to the evaluation engine are presented in Table 26 and Table 27. Overall 46,030 pairs of pedigree records were compared during the evaluation process, out of which 1,302 are within-message pairs and 44,728 are between-message pairs. It can be seen that the pedigree record counts in type restricted evaluation are smaller than those in the unrestricted evaluation, which results in higher Precision, Recall, and F-score, as expected.

Table 26: Evaluation statistics for sequential GA^N

Evaluation Mode	Correctly Associated	Incorrectly Associated	Incorrectly Not Associated
Unrestricted	29,492	2,682	12,554
Type Restricted	29,349	2,382	8,836

Table 27: Evaluation statistics by entity type for sequential GA^N

Type	Correctly Associated	Incorrectly Associated	Incorrectly Not Associated
Person	27,736	1,941	7,612
Location	321	164	442
Organization	1,351	561	1,080
Vehicle	8	0	6
Un-typed	77	16	3,414

The computational results for data association are presented in Table 28. The Precision, Recall and F-score in the table represent the accuracy of the association (type restricted). Higher values typically indicate greater accuracy. These metrics are likely to improve in the future, as the hard data processing techniques mature, providing richer information for hard+soft and hard+hard association.

Table 28: Computational results for data association

No.	Procedure	Precision	Recall	F-Score	Computing Time (seconds)
1	GA^N (Sequential)	0.925	0.769	0.839	794
2	MDADC (MR)	0.938	0.772	0.847	65
3	CPP (Streaming)	0.915	0.796	0.851	1,312 (10 s / graph update)

Since the GA^N formulation is tighter than MDADC and the MDADC formulation is tighter than CPP, the accuracy of GA^N should have been greater than MDADC and the accuracy of MDADC should have been greater than CPP. However, the results obtained from the computational experiments do not seem to follow this reasoning. The reason behind these counter-intuitive results can be explained as follows. As mentioned before, the within message

co-referencing is performed by Tractor. Therefore the main assumption of GA^N and MDADC models is that there are no duplicate references within a particular message. If Tractor were to miss any of the within message co-references, then this imprecision is propagated in the data association results. The CPP formulation does not assume the absence of duplicate references within a particular message. Therefore the CPP formulation associates more node pairs as compared to MDADC and GA^N, which makes its Recall the highest. The MDADC formulation does not have the complicating edge-association constraints, which could prevent some nodes from being associated, making its Recall the second highest. The most constrained GA^N formulation takes the third place. If the input data to the GA^N is clean (no ambiguity in similarity scores and no duplicate references within same message), then it will likely outperform the MDADC and CPP formulations in terms of accuracy.

The sequential Lagrangian procedure for GA^N formulation takes the second longest time to solve due to the complexity of the model. The Map/Reduce Lagrangian procedure for MDADC requires much less time to solve, because multiple processors share the computational burden. If the data size is large, the sequential Lagrangian heuristic for GA^N will prove to be a bottleneck. On the other hand, MDADC formulation solved using Map/Reduce will provide a quick and reasonably accurate solution and it can be easily applied to large sized problems given the necessary hardware. The cumulative time required for Streaming Entity Resolution algorithm, is the longest. However it translates into an average of 10 seconds per graph update, which is better than re-solving the data association problem on the entire dataset using one of the batch algorithms.

3.2.1.10 Hard+Soft Data Association References

[H.1] K. Date, G. A. Gross, R. Nagi. 2014. Test and Evaluation of Data Association Algorithms in Hard+Soft Data Fusion. Proceedings of the 17th International Conference on Information Fusion (Fusion 2014).

[H.2] M. Prentice, M. Kandefer, and S. Shapiro, "Tractor: A Framework for Soft Information Fusion," Proceedings of the 13th International Conference on Information Fusion (Fusion 2010), 2010.

[H.3] Levenshtein, V. I. (1965), "Binary codes capable of correcting deletions, insertions, and reversals." Doklady Akademii Nauk SSSR 163 (4): 845–848.

[H.4] Princeton University "About WordNet." WordNet. Princeton University. 2010. http://wordnet.princeton.edu.

[H.5] M. Jenkins, G. Gross, A. Bisantz, R. Nagi, "Towards Context Aware Data Fusion: Modeling and Integration of Situationally Qualified Human Observations into a Fusion Process for Intelligence Analysis," Journal of Information Fusion, (Accepted June 2013).

[H.6] M.P. Jenkins, G. A. Gross., A. M. Bisantz, and R. Nagi, "Towards context-aware hard/soft information fusion: Incorporation of situationally qualified human observations into a fusion process for intelligence analysis." In Proc. of the 2011 IEEE First International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), Feb. 22-24, Miami Beach, FL, pp. 74-81.

[H.7] Luiz Antonio Lorena and Marcelo Narciso. Relaxation heuristics for a generalized assignment problem. European Journal of Operations Research, 91(3):300-310, 1996.

[H.8] G. Tauer, R. Nagi, and M. Sudit, “The graph association problem: mathematical models and a Lagrangian heuristic”. Naval Research Logistics, vol. 60, pp. 251-268, April 2013.

[H.9] G. Tauer, and R. Nagi, “A Map-Reduce Lagrangian heuristic for multidimensional assignment problems with decomposable costs,” Parallel Computing, 39(11), pp. 653-658, November 2013.

[H.10] H. Bandelt, A. Maas, and F. Spieksma. Local search heuristics for multi-index assignment problems with decomposable costs. Journal of Operational Research Society, 55:694–704, 2004.

[H.11] G. Tauer, K. Date, R. Nagi, and M. Sudit, “An incremental graph-partitioning algorithm for entity resolution,” Under Revision. To be submitted to Transactions on Knowledge Discovery from Data.

[H.12] Grötschel M.; Y. Wakabayashi, “A Cutting Plane Algorithm for a Clustering Problem”, Mathematical Programming 45: 59-96, 1989.

[H.13] Dorndorf U.; E. Pesch, “Fast Clustering Algorithms”, ORSA Journal on Computing, 6: 141-168, 1994.

3.3 Pennsylvania State University

3.3.1 Pennsylvania State Abstract

This report summarizes the activities performed by the Pennsylvania State University (PSU) during the fifth year of this MURI project in support of the University at Buffalo (under the direction of Dr. Rakesh Nagi and Dr. Moises Sudit). The Penn State activities during this period focused in five areas; (1) test and evaluation, (2) evolution of the cyber infrastructure for distributed hard and soft data fusion, (3) enhanced hard sensor processing, (4) automated sense-making algorithms and, (5) visual analytics and cognitive assessment. The remainder of this report provides a summary of accomplishments, project statistics and summary, publications and identification of key personnel. Section 3.3.4 provides an overall summary of the five year accomplishments of the Penn State team. Section 3.3.8 provides additional details on hard sensor processing techniques. Finally, additional details on the fifth year accomplishments are available in the accompanying papers and documents.

3.3.2 List of Papers Submitted or Published

- Papers published in peer-reviewed conference proceedings

35. J. Rimland and M. Ballora, (2014), “Using Complex Event Processing (CEP) and vocal synthesis techniques to improve comprehension of sonified human-centric data,” *Proceedings of SPIE 2014*, Baltimore, MD, May 6, 2014.

36. J. Rimland and M. Ballora, (2014), “Using vocal-based sounds to represent sentiment in complex event processing”, *Proceedings of the International Conference on Auditory Display (ICAD),* 2014

37. J. Rimland, S. Shaffer and D. L. Hall, (2014) "A hitchhiker's guide to distributed hard and soft information fusion infrastructure development", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain.
38. S. Shaffer, (2014), "Automatic theory generation from analyst text files using coherence networks", *Proceedings of the SPIE 2014 Conference, Baltimore*, MD, May 6, 2014.
39. G. Cai and J. Graham (2014), "Semantic data fusion through visually-enabled analytical reasoning", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain
40. G. Cai, G. Gross, J. Llinas and D. Hall (2014), "A visual analytic framework for data fusion in investigative intelligence", *Proceedings of the SPIE volume 9122, Next-Generation Analyst II*. Baltimore, Maryland, USA
41. John P. Morgan and Richard L. Tutwiler, " Real-Time reconstruction of depth sequences using signed distance functions ", *Proceedings of the SPIE 2014 Conference*, Baltimore, MD, May 6, 2014

f) Presentations

- Papers presented at peer-reviewed conferences

16. J. Rimland and M. Ballora, (2014), "Using Complex Event Processing (CEP) and vocal synthesis techniques to improve comprehension of sonified human-centric data," *Proceedings of SPIE 2014*, Baltimore, MD, May 6, 2014.
17. J. Rimland and M. Ballora, (2014), "Using vocal-based sounds to represent sentiment in complex event processing", *Proceedings of the International Conference on Auditory Display (ICAD),* 2014
18. J. Rimland, S. Shaffer and D. L. Hall, (2014) "A hitchhiker's guide to distributed hard and soft information fusion infrastructure development", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain.
19. S. Shaffer, (2014), "Automatic theory generation from analyst text files using coherence networks", *Proceedings of the SPIE 2014 Conference, Baltimore*, MD, May 6, 2014.
20. G. Cai and J. Graham (2014), "Semantic data fusion through visually-enabled analytical reasoning", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain
21. G. Cai, G. Gross, J. Llinas and D. Hall (2014), "A visual analytic framework for data fusion in investigative intelligence", *Proceedings of the SPIE volume 9122, Next-Generation Analyst II*. Baltimore, Maryland, USA

22. John P. Morgan, Richard L. Tutwiler, " Real-Time reconstruction of depth sequences using signed distance functions ", *Proceedings of the SPIE 2014 Conference*, Baltimore, MD, May 6, 2014

- Other presentations

1. J. Graham, "SYNCOIN: a synthetic dataset for evaluating hard and soft fusion algorithms," presentation to SI Org University Innovation Day Share [IT], 2 August 2012, Chantilly, VA

g) Manuscripts

5. J. Graham et al (2014), *Analyst Workbench Instructional Guide*, Technical report for the NC2IF Research Center, August, 2014 (30 pages)

h) Books and Book Chapters

6. D. Hall, J. Llinas, C. Chong, K. C. Chang, editors, *Distributed Data Fusion for Network-Centric Operations,* CRC Press*,* August, 2012
7. D. L. Hall, "Perspectives on Distributed Data Fusion", chapter 1 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press*,* August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang
8. J. Rimland, "Service-Oriented Architecture for Human-Centric Information Fusion," chapter 13 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press*,* August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang
9. D. Hall, "The Emergence of Human-Centric Information Fusion," chapter 27 in *Distributed Sensor Networks*, 2nd edition, 2012, edited by S. Iyengar and R. Brooks

i) Theses and dissertations

7. S. R. Nimmala (2014), *Architectural considerations for context aware applications in mobile cloud computing environment*, M.S. thesis in Computer Science and Engineering, The Pennsylvania State University, August 2014
8. J. C. Rimland, (2013), *Hybrid human-computing distributed sense-making: Extending the SOA paradigm for dynamic adjudication and optimization of human and computer roles*, Ph.D. dissertation in Information Sciences and Technology, The Pennsylvania State University, August, 2013

Honors and Awards – N/A

Titles of Patents disclosed during the reporting period – N/A

Patents awarded during the reporting period – N/A

Graduate Students

Name	Per Cent Supported
Dong Chen	50%
Rob Grace	50%
Spoorthi Rao Nimmala	25%
Na Sun	25%
Jeff Rimland	50%
Total Number:	5

Post Doctorates

Name	Per Cent Supported
Jeff Rimland[6]	25%
Total Number:	1

Faculty

Name	Per Cent Supported
D. Hall[7]	0.0 %
M. McNeese	2.5%

6 During this fifth year, Jeff Rimland transitioned from a graduate student to a post-PhD/Research Associate

7 D. Hall's time is funded by the Pennsylvania State University without charge to the project – amount of time dedicated to the project is 10 %

J. Graham	10.0%
R. Tutwiler	10.0%
G. Cai	12.0%
S. Shafer	10.0%
Total Number:	6

Under Graduate Students

Name	Per Cent Supported
Emily Catherman	15%
Total Number:	**1**

Student Metrics

The number of post-graduates & PhDs funded during this period[8]	1
The number of under-graduates funded during this period	1
The number of undergraduates funded who graduated during this period	1
The number of undergraduates who graduated during this period with a degree in science, mathematics, engineering, or technology fields	1
The number of undergrads who graduated during this period and will continue to pursue a graduate or PhD degree in science, mathematics, engineering or technology fields	0
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0	1
Number of graduating undergrads funded by a DoD funded Center of Excellence grant for Education, Research and Engineering	1
The number of undergrads who graduated during this period and intend to work for the Department of Defense	1
The number of undergraduates who graduated during this period and will receive scholarships	0

[8] During the fifth year, Jeffrey Rimland transitioned from graduate student status to Post-PhD/research associate status

or fellowships to further studies in science, mathematics, engineering or technology fields	

Masters Degrees Awarded (1)

- S. R. Nimmala (2014), *Architectural considerations for context aware applications in mobile cloud computing environment*, M.S. thesis in Computer Science and Engineering, The Pennsylvania State University, August 2014

PhDs Awarded (1)

- J. C. Rimland, (2013), *Hybrid human-computing distributed sense-making: Extending the SOA paradigm for dynamic adjudication and optimization of human and computer roles*, Ph.D. dissertation in Information Sciences and Technology, The Pennsylvania State University, August, 2013

Other Research Staff – None

Technology transfer

- Continued collaborations with Penn State Police Services regarding stadium protection/campus activities
- Collaboration with Raytheon Corporation on a related IR&D project
- Development of information fusion concepts with the SI Organization
- Lockheed Martin (planning related IR&D project)
- Discussions with USAF NORTHCOM regarding Homeland Security
- Discussions with MIT Lincoln Laboratory regarding test and evaluation and human analyst in the loop concepts
- Meetings and visit to the Naval Surface Warfare Center (Crane Division) regarding hard and soft fusion and test and evaluation, including access to GBOSS equipment and software
- Meeting with DHS/Transportation Security Administration (TSA)
- Discussions with the Boeing Corporation regarding collaboration
- Continued interaction with the Pennsylvania State University Applied Research Laboratory

- The SYNCOIN data set was shared with the following organizations and individuals (during both year 3 and year 4) with continued follow up during the 5th year for SYNCOIN updates:
  - Peter Willet, University of Connecticut
  - Gavin Powell, ADS Innovation Works, UK, government technical area lead for TA 6 - Distributed Coalition Information Processing for Decision-Making
  - David Nicholson, BAE Systems, London, UK
  - David Dearing, Stottler Henke Associates
  - David Braines, Hursley Emerging Technology Services
  - Erick Blasch, Air Force Research Laboratory Sensors Directorate (AFRL/SNAA)
  - Marco Pravia, BAE Systems
  - Kamal Premaratne, University of Miami
  - James Law, SPAWARSYSCEN – U. S. Navy Space and Naval Warfare Systems Center
  - Chase Cotton, Network Science Collaborative Technology Alliance Program (CTA), U. S. Army Research Laboratory
  - ETURWG – Evaluation of Techniques for Uncertainty Representation Working Group, International Society of Information Fusion (ISIF)
  - International Technology Alliance
  - Brian Simpson, Raytheon Corporation
  - Simon Maskell, QinetiQ, UK
  - Charlotte Shabarkh, Aptima, Woburn, MA
  - Brian Ulicny, VIStology, INC, Framingham, MA
  - Dr. Joan Carter, Institute for Defense Analysis, Alexandria, VA
  - Network Science Collaborative Technology Alliance, University of Illinois, Champaign, IL
  - Jim Fleming, Saffron Technology, Cary, NC
  - Charles Morefield, Arctan, Arlington, VA
  - Rick Beckett, Overwatch, Textron, Philadelphia, PA
  - Dr. Tony Penza, MIT Lincoln Laboratory
  - Naval Surface Warfare Center, Crane Division

- Organized and hosted technical sessions at national conferences
  - Next Generation Analyst II: Special one-day session organized for the SPIE Conference on Sensing Technology and Applications, May 2013 in Baltimore, MD
  - International Society of Information Fusion - FUSION 2014 Conference held in Salamanca, Spain, July, 2014 – organized a special session on advances in hard and soft information fusion

3.3.3 Pennsylvania State Accomplishments and Narratives of Research Efforts

During this reporting period, advances were made in five major areas; (1) test and evaluation, (2) cyber infrastructure, (3) hard data processing, (4) automated sense-making algorithms, (5) visual analytics and cognitive assessment. These are summarized below and described in additional detail in referenced papers.

1) *Test and evaluation* – The SYNCOIN data set continued to be enhanced to meet the evolving needs for test and evaluation of hard and soft fusion algorithms and for demonstration purposes [I.1.2]. In addition, the Penn State team participated in meetings of the International Society of Information Fusion Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG). In addition, effort continued on the development and refinement of TML representations for the joint 5th year demonstration. Data were updated to include entity characteristic and identity meta-data to support association and correlation as input to University at Buffalo graph-matching techniques.

2) *Cyber infrastructure* – Work continued on the evaluation and implementation of emerging standards and tools to facilitate distributed hard and soft data fusion. An extensive tutorial document was created [I.2.3]. In addition, a new framework (extending the framework of [I.1.24]) for distributed context-aware sensing, fusion and decision-making was developed and demonstrated for a cloud computing environment [I.5.1].

3) *Hard data processing* – Work continued on the automation of the target identification and characterization of the hard sensor data (viz., collected as part of the vignettes intended for inclusion in the SYNCOIN data set). We completed the automated classification of human forms from 2-D and 3-D map fused data. A multi-level association technique was designed and implemented to translate parametric data to provide “story book” scene characterization. A final data collect using KINECT was performed to merge the color tracker and depth map tracker [I.2.7].

4) *Automated sense-making algorithms* – Complex Event Processing (CEP) methods and Multi-Agent Systems (MAS) were developed and applied to the SYNCOIN data [I.2.3]. A new coherence network algorithm was implemented and demonstrated for SYNCOIN data [I.2.4]. We linked Complex Event Processing/Multi-Agent Systems and Coherence Network processing to support hypothesis generation and analyst focus of attention [I.2.3].

5) *Visual analytics and cognitive assessment* – Effort continued on the development of a “data analytics/analysis visualization tool kit”. A web based visualization system has been developed to allow display (and interaction among) multiple panels providing a geographical map, a data/”event window”, social network displays, timeline view, and workspace for hypothesis generation and analysis [I.2.5][I.2.6]. Activities included; i) extending the cognitive task analysis to improve understanding of analytical reasoning processes, ii) extending the current toolkit (e.g., text analysis and recommender support)

for enhanced analysis capabilities, iii) investigation of techniques to enable collaborative analysis with multiple distributed analysts, iv) integration of visual interaction with computer automated processing, v) implementation of a multi-view data exploratory analysis (data dashboard) with several coordination strategies, vi) development of a novel sonification techniques using CEP and vocal synthesis [I.2.1], vii) integration of visual data fusion with intelligence analysis, and viii) fielding of a beta version of the toolkit and initiation of human in the loop experiments.

The following sections review the overall hard and soft data fusion processing concepts with an emphasis on the activities performed by Penn State researchers and provide additional insight regarding the SYNCOIN data set, the data visualization and analytics toolkit, and automated sense-making algorithms. Additional details on the hard sensor processing are provided in Section 3.3.8.

Distributed Hard and Soft Data Processing Concepts

The overall concept for processing hard and soft data has evolved during this five year effort. The original concept is described in the proposal and detailed in [I.1.10], and in [I.2.12]. The concept involved separate processing and fusion of hard (physical sensor data) and soft (human observed data) followed by centralized fusion of the hard processing results and the soft processing results. The basic concept has remained unchanged, although details of the processing flow have evolved with increasing understanding of the basic functions required for each processing sub-process. An overview of a current view is shown in Figure 40 with detail provided for the University at Buffalo (UB) processing. The dashed line in the center of the figure represents yet to be provided detail about the Penn State processing flow. The conceptual processing flows in the bottom part of the figure represent planned processing by Dr. Amir Shirkhodaie of the Tennessee State University and by Dr. Corso of the University at Buffalo, respectively.

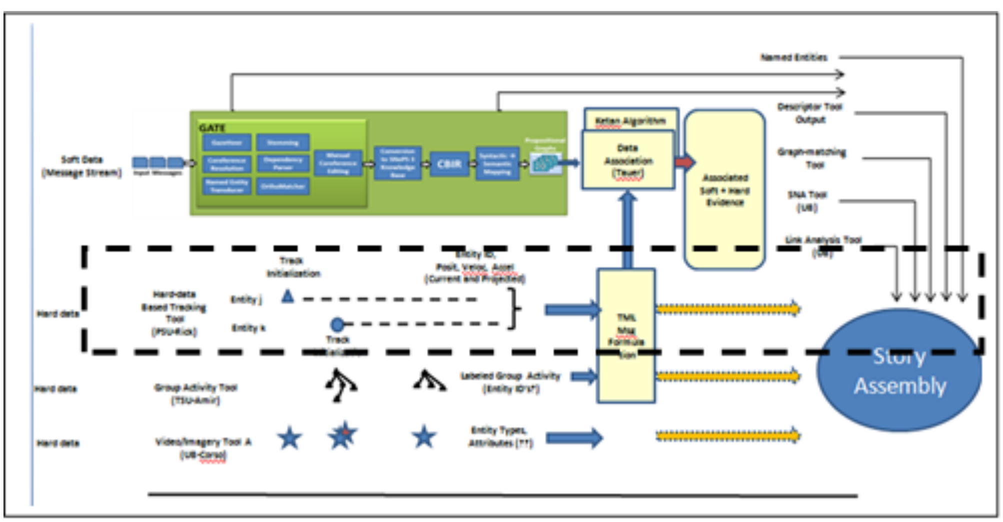


Figure 40: Conceptual Framework for Hard/Soft Fusion Processing [I.1.9]

The current processing concept for the Penn State functions is shown in Figure 41. This represents an expansion of the dashed rectangle shown in the middle part of Figure 40 above.

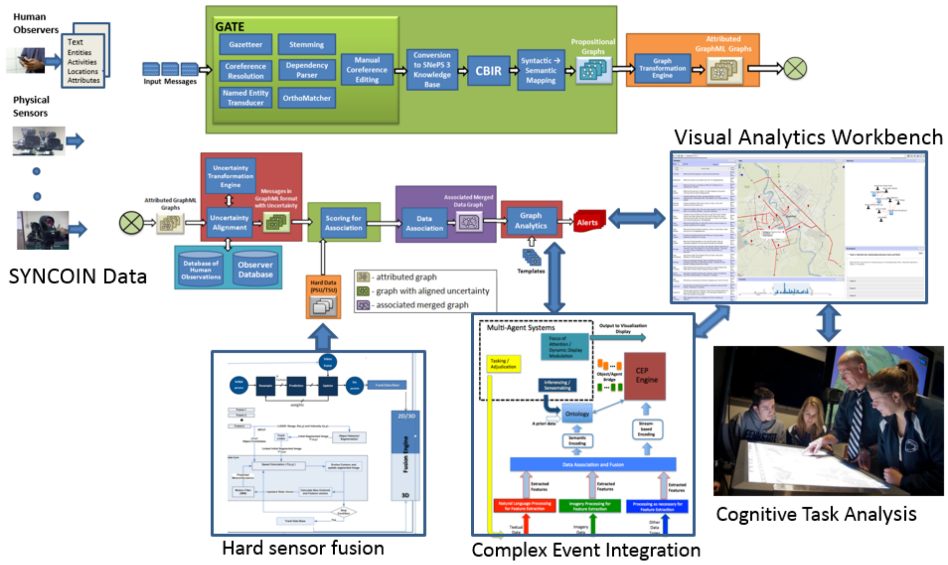


Figure 41: Penn State components of the overall hard and soft fusion processing flow

We note that the processing shown in Figure 41 includes the following:

i) Detailed processing flow for fusion of 2D and 3D hard sensor data using multiple hypothesis filters;

ii) Computer and human generated meta-data to add to the fused hard sensor tracking information;

iii) Processing of soft sensor data via addition of semantic meta-data based on both semi-automated and human generated data;

iv) Creation of an accessible meta-data base which includes both original soft sensor data viz., the SYNCOIN data [I.2.14], [I.2.18], [I.2.10], [I.4.10] along with enhanced meta-data [I.1.7], [I.1.6], [I.4.11], as well as the original hard sensor data (viz., collected via the deployment of physical sensors at the Pleasant Gap, PA fire safety training facility (https://sss.cpi.edu/est and [I.2.15]);

v) Focus of attention processing via Complex Event processing (CEP) [I.5.7] and coherence network processing [I.2.3]; and finally

vi) Interaction with one or more analysts via a geographical information display augmented with data overlays, access to hard and soft data, display of evolving processing and analysis results, along with access to traditional analyst tools such as Analyst's Notebook [I.4.13].

Details of the human computer interaction and operational concepts for how an analyst adjudicates between requests for information (via PIRs, RFIs, etc.), the cognitive and analysis processes for the analyst(s) are still evolving [I.2.26] [I.4.13]. [I.5.1] explored architectural concepts for context-aware applications in a mobile cloud computing environment. She extended the framework, originally developed by [I.1.24] to explicitly consider the tradeoff between functions performed on a mobile sensing/reporting device (a smart phone) and cloud-based computational resources. She demonstrated the framework for a smart phone travel application.

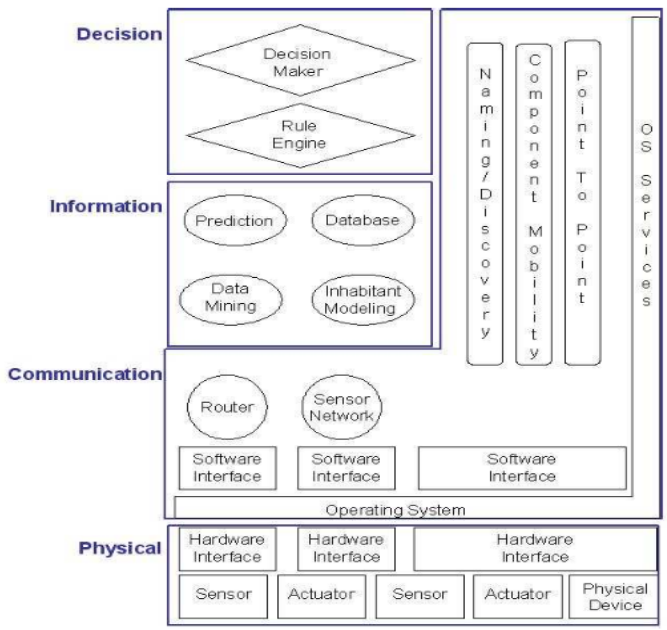


Figure 42: General framework for context-aware computing in a cloud-based environment [I.5.1]

Additional details regarding the hard sensor processing flow and fusion is provided in Figure 43 (adapted from [I.2.33]).

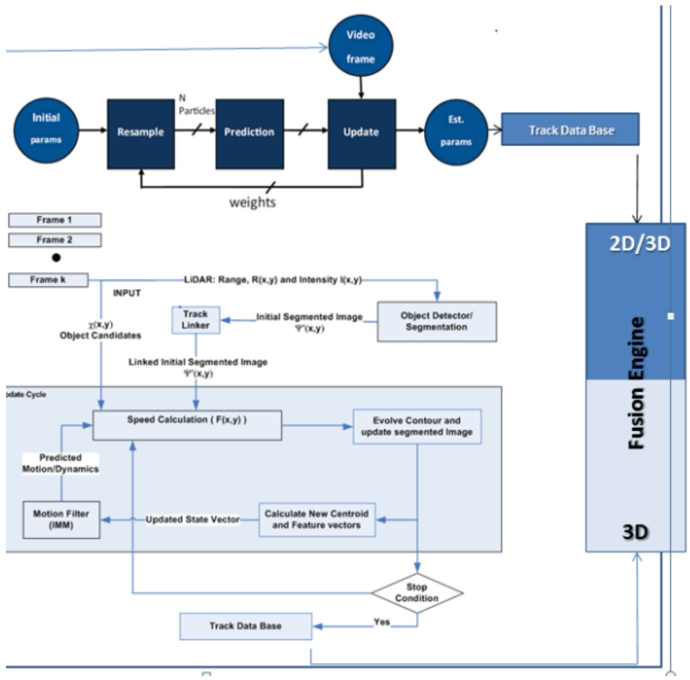


Figure 43: Enlarged View of Hard Sensor Processing Flow

Test and Evaluation: Continued evolution of SYNCOIN data[9]

By the end of year two, we had completed the initial development of a synthetic coin inspired data set (SYNCOIN) to support the test and evaluation of emerging hard and soft data fusion algorithms and techniques. The data set is inspired by a Counter Insurgency (COIN) scenario in Bagdad. The data includes over 600 messages ("soft" data) and synthetic complimentary hard data (e.g., simulated physical sensor data). The scenarios cover a four month period: 1 January – 10 May 2010, centered in Baghdad, Iraq. The central theme throughout the dataset involves Improvised Explosive Device (IED) operations and associated networks. Several sub-plots or threads are woven throughout the message set – all dealing in some measure with the people, motivations and intent of IED related activities. Specific care was given to NOT emulate actual IED tactics, counter-tactics or operational tradecraft; hence U.S. unit designators and agency names were largely omitted. Overviews of this synthetic data set are provided by [I.2.18], and [I.2.26]. The overall test and evaluation concept for this MURI project is described by [I.2.10].

The SYNCOIN scenario emulates many of the complexities and challenges incumbent of COIN operations in Iraq without disclosing specific collection strategies, methods or means. The message set deliberately down-plays contentious aspects of counter-insurgency operations such as interrogations and the targeting of humans for elimination. The foundation of the message set is the reporting of "soft" data; i.e., information collected by humans on human activities; however, it also represents multiple "hard" data opportunities; i.e., reports that reflect the

[9] We provide here information on the evolution of the SYNCOIN data to assist the reader in understanding the significance and role of this extensive data set.

collaboration of soft reports with hard sensor means. The intent has been to create synthetic hard data products to emulate the type of analysis products that would accompany respective soft reports. The scenario involves a hypothetical situation in Bagdad in 2010, a period in transition. The SYCOIN messages simulate brief summations of event reports, observations, findings and analysis of COIN-related activities from a street-level view. Higher-level observations are also presented to represent agency or headquarters' (HQ) views. The target audience of the message set is the battalion commander

During the third year of this project detailed products for each of these threads were created to assist the test and evaluation process for fusion algorithms. In particular, the "ground truth" products included the following.

- A listing of all SYNCOIN synthetic messages identified by vignette/threads [I.4.10];
- A textual "scene setter" for the overall SYNCOIN messages and for each vignette/thread [I.1.2];
- Description of the build strategy [I.4.7];
- An acronym list [I.1.3];
- Identification and location of all events and activities - providing both latitude, longitude, MILGRID coordinates and associated labels of places, events and activities [I.1.4];
- Reference maps for SYNCOIN [I.1.5];
- Database schema for each thread (events, objects, locations, persons, and activities) [I.1.6];
- *Analyst Notebook* social network analysis diagrams for each thread [I.1.7]; and
- Word Cloud diagrams (based on *Wordle*) for each SYNCOIN thread [I.4.11].

During the fourth year of the project, work continued to evolve and refine the SYNCOIN data set. In particular,

- Application of SYNCOIN data sets into AXIS Pro visual analytic software created enhanced ground truth products.
- SYNCOIN/AXIS Pro integration was used to develop Analyst Training protocols for the creation of geo-spatial visualization products.
- A back-end product of SYNCOIN/AXIS Pro integration was a comprehensive set of descriptive meta-data with unique hexadecimal identifiers for each SYNCOIN entity.
- Entity matching using unique entity identifiers reduced overall data ambiguity by matching known names with alias, or alternate spellings, etc.
- AXIS Pro/SYNCOIN integration process facilitated a comprehensive review of the analytic process for conducting human-centric analysis and sense making of disparate data.
- Refinements were made to aid the Test and Evaluation process, in particular the designation and/or refinement of geo-reference data for key named events.

- Additional SYNCOIN messages were created to facilitate the creation/description of micro-vignettes that became the focus point of the hard sensor processing effort.

During year 5, we continued the evolution of the SYNCOIN data set including addition of new physical sensor data, new soft and hard sensor data links, and creation of new meta-data associated with the hard sensor data. We continued the dissemination of the SYNCOIN data to the data fusion community and participated in discussions and planning of the ETURWG – Evaluation of Techniques for Uncertainty Representation Working Group, International Society of Information Fusion (ISIF).

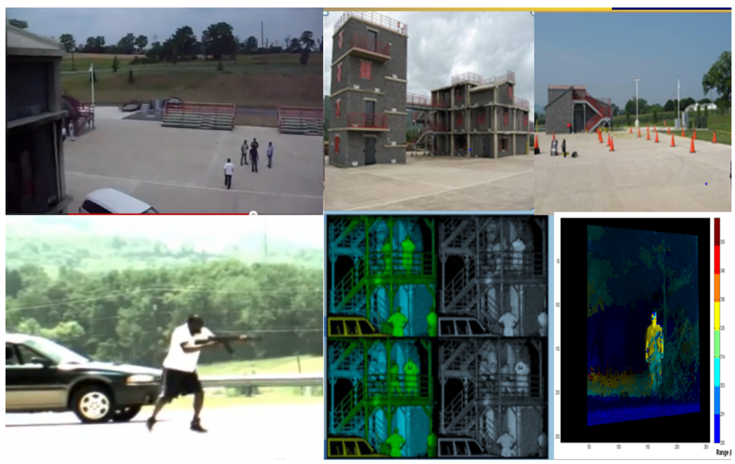


Figure 44: Collage of images from the Pleasant Gap facility

Hard Data Processing

During the fourth year, research continued on processing data collected during the 3rd year of the project. Recall that at the end of the 3rd year, data collections were conducted at the Centre County Public Safety Training facility in Pleasant Gap, Pennsylvania [I.1.9]. The fire safety facility serves as an excellent location for the hard sensor experiments involving humans in the loop. Figure 44 shows a collage of images from the Pleasant Gap facility showing the overview of the three buildings and open area (in the upper part of the figure), examples of student "actors" portraying aggressor actions, and sample Lidar images in lower part of the figure. Additional data were collected during the fifth year of this project.

The data shoot was conducted on July 23- 25th, 2012, and involved an urban setting to augment the demonstration planning and environment. Multiple "micro-vignettes" were developed to allow creation of multiple hard data injections into SYNCOIN. Scripted scenarios involved humans interacting in a building, multiple vehicles, simulated crowd (market-place) activities, pickup of a man, assault activities, IED events, etc. Sensors included; Lidar, VNIR four camera surveillance suite, and VNIR HD gen-locked stereo camera pair. This hard sensor data was merged into the SYNCOIN message threads (e.g. the SYNCOIN threads were augmented to provide a motivation and link to the collected Pleasant Gap data (which was "repositioned" in time and location to the appropriate Bagdad locations commensurate with the message threads).

Processing of the collected data continued with the development and refinement of algorithms for target tracking and characterization/identification. In addition, during year 5, we collected additional KINECT data to assist experiments with multi-sensory data level fusion and automated target characterization and identification. Details of the hard sensor processing are provided in Section 3.3.8.

Generation of TML for Joint Demonstrations

In order to support the test and evaluation of algorithms developed across the MURI team, Penn State developed TML files for the collected hard sensor data described above. Figure 45 shows an example of a single frame of hard sensor data and identifies six objects (including a vehicle and different people). For each of these objects, the automated hard sensor processing described previously provides information about the object location, type or characteristic, observation time, and related data.

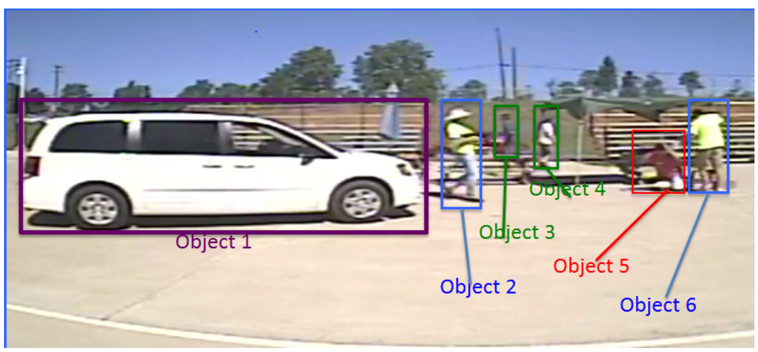


Figure 45: Sample snapshot frame from hard sensor data

TML formats were developed and specified, and TML data were populated for each entity, location and observation time.

In year 5, a web service-based approach was applied to the conversion of raw sensor "track" files into TML documents. This process, illustrated in Figure 46, relies on multiple Advanced Message Queuing Protocol (AMQP) queues, worker daemons for the actual conversion and formatting of TML, and endpoint applications that divide incoming track files into single line segments, and reconstruct single lines into complete TML files.

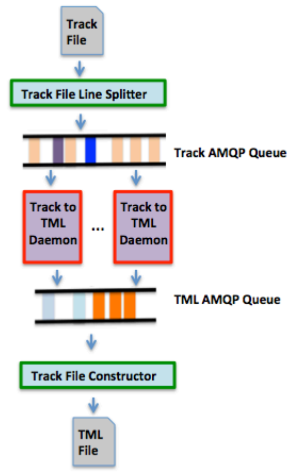


Figure 46: Architecture for allowing massively parallel conversion of very large track files into TML

This approach facilitated the conversion of very large track files into TML. By dividing the track files on a line-by-line basis and uploading individual lines into an AMQP queue for processing, multiple Track-to-TML daemons can operate on a single large file in parallel.

Data Visualization Toolkit

In order to support the development of end-to-end processing and analysis of hard and soft fusion, a data visualization and analysis toolkit was developed using web-based tools. The purpose of the toolkit is to allow researchers to rapidly explore the utility of emerging cognitive aids and automated processing algorithms for improved situation awareness. The tool supports both data-driven analysis as well as analyst initiated hypothesis-driven analysis. The perspective is that, much like human situation awareness of the world in which simultaneous input is received from human senses and "alerts" and human attention is focused by cognitive intent. This extends the first three years of this MURI effort by explicitly considering the "analyst in the loop" for continuous inference processing.

A sample of the user interface for the tool is shown in Figure 47. The figure illustrates that multiple "views" of the evolving situation and associated data are presented simultaneously. A central geographical display anchors the analysis, allowing analysts to show, and interact with, data from a geo-spatial viewpoint. On the left hand side of the display, message data are shown – listing the individual messages (obtained from human reports or generated by automated hard sensor processing operations). The bottom part of the figure shows a timeline and a histogram of the amount of message data received per unit time. Finally on the right hand side of the display is shown a graphical display of social network data.

These display inserts are dynamically linked. For example, if a user conducts a query (based on a time window, geographical region, and content information), those messages would be displayed on the left hand insert. Simultaneously, the data would be automatically displayed on the map, a social network display would be created indicating links and associations between

named entities on the right hand side of the display, and finally the data timeline would be populated at the bottom of the display. The data are individually linked as well. Thus, if an analyst selects a data point shown on the map, the associated message from which the data were derived would be highlighted as well as the pertinent nodes in the social network display. This interactive toolkit allows an analyst to formulate and explore emerging hypotheses about events, activities, and entities, readily moving across the key questions of "who, what", "where", "when" and "why".

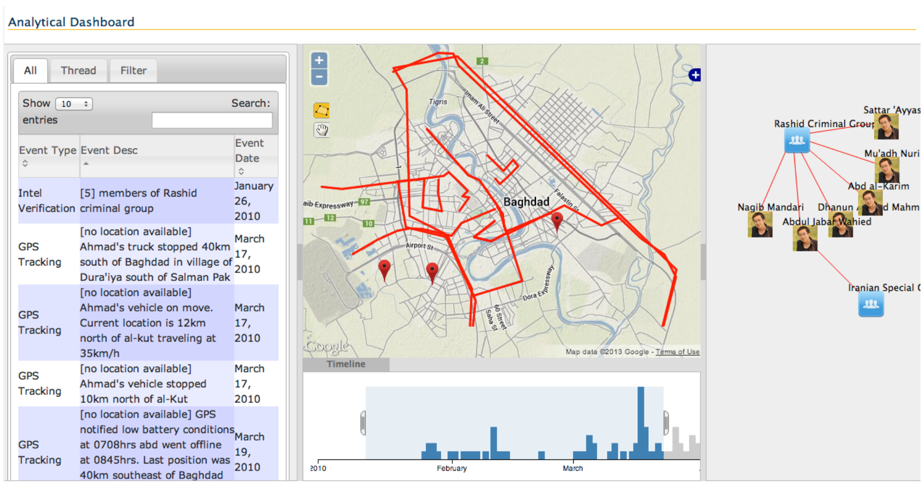


Figure 47: Screen shot of visualization/analysis toolkit

During the 5th year the toolkit was extended to incorporate analyst interaction in the creation and assessment of hypotheses. Feedback from users was obtained to refine the interaction and availability of analysis tools. An *Analyst Workbench Instructional Guide* has been developed [I.4.13] to assist in helping new users in understanding the utility of the toolkit.

Analysis Concept/Cognitive Task Analysis

The development of the visualization/analysis toolkit proceeded in parallel with the development of analysis concepts and cognitive task analysis. Under the leadership of Col. Jake Graham, a team of students conducted "guided" analysis of the SYNCOIN data acting as if they were military analysts supporting a command function. The team of students receives general guidelines and directions via "daily" requests for information and evolving understanding of indications and warnings (I&W). As the student team conducts their analysis, receiving input data (via SYNCOIN messages and related hard sensor data), develop hypotheses regarding activities, events, and situations, meta-cognition analysis is performed. The processes and mental flow of the individual and team cognition was assessed to determine what tools and cognitive aids might be useful to assist the process. This analysis guided the refinement of the toolkit and extension to support new cognitive aids.

Automated Inference Tools

During the fourth year effort, we investigated an improved information framework for human in the loop processing (including viewing humans as observers to augment hard sensor data as well as humans collaborating with automated inference processes. In particular, two types of algorithms were investigated; i) complex event processing and ii) multi-agent systems. This processing concept it illustrated in Figure 48. We term this a hybrid, human-computing distributed sense-making (HHCDSM) architecture. During year 5, these algorithms were refined and applied to the SYNCOIN data set.

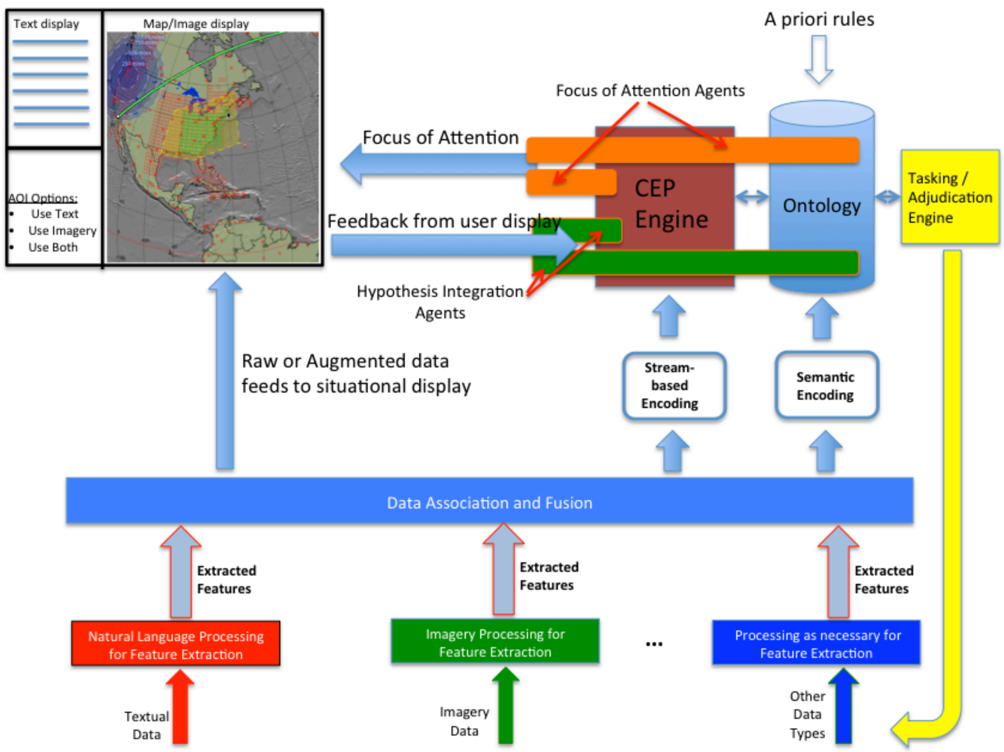


Figure 48: Information architecture for HHCDSM showing distributed heterogeneous data, complex event processing, and the use of a multi-agent systems approach for focusing user attention and integrating top-down and bottom-up processing

In this architecture, special consideration is given to top-down (i.e. hypothesis driven) vs. bottom-up (i.e. data driven) processing and inferencing. The Complex Event Processing (CEP) field (summarized below) provides excellent facilities for rule-based sense-making over broadly distributed and heterogeneous input streams. However, stream-based processing is largely restricted to a data-driven approach. Agent-based systems, on the other hand, are well suited to hypothesis-driven analysis of data. The HHCDSM approach embeds mobile agent "shells" into a high performance stream-based CEP engine in order to introduce the capability to perform hypothesis-based reasoning, response rapidly to user inputs, and otherwise modulate the activity of the CEP engine in light of changing hypotheses (see Figure 48). A brief summary of CEP and MAS are provided below.

Complex Event Processing

Complex event processing (CEP) addresses the challenge of combining multiple heterogeneous data streams into a hierarchical structure that can represent higher order events

and semantic meaning through the application of rules and filters at multiple levels of information [I.1.11]. For example, the three individual events of a man wearing a tuxedo, a woman wearing a gown, and people throwing rice can be combined into the single event of "a wedding" with a certain probability. Then individual "wedding" events could be combined with other events at that level to determine higher-level trends. Complex event processing is often used in conjunction with ontological representations of data to facilitate the transformation from stream-based data into organized event hierarchies [I.1.12].

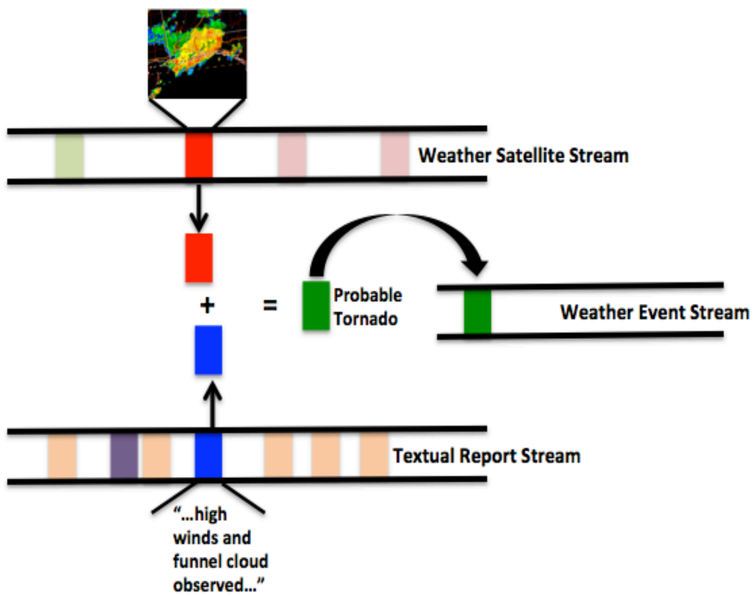


Figure 49: Complex Event Processing (CEP) uses rules and filters to combine and aggregate events

In Figure 49, the red and blue blocks represent low-level events extracted from streams of physical sensor data and human-centric reports (respectively). Note that Figure 49 refers to an example application involving monitoring a potentially threatening weather condition. The figure shows the red and blue blocks being combined (typically through a process of rule-based aggregation) into the green block, which is demarcated as a "probable tornado" event, and then placed into an additional data stream. This higher level event could be combined with other similar-level events (e.g. "probable landslide" or "high winds") and further aggregated into an even higher level event or trend (e.g. "there has been an increase in severe weather in July"). This process can repeat in a fractal manner until a rich hierarchy of events is represented.

Although the CEP formalism has primarily arisen out of financial and stock trading applications, it has recently been applied to other applications including smart energy [I.1.13], heterogeneous sensor network processing [I.1.12], radio frequency identification (RFID) middleware [I.1.14], and data fusion applications for "strategic intelligence" [I.1.15]. In the latter, [I.1.15] use CEP to perform Joint Directors of Laboratories (JDL) data fusion process model Level 2 and 3 tasks of situation and threat assessment via the CEP tasks of *filtering*,

aggregating, and *detection* via event pattern rules (EPRs). These works establish CEP as an emerging area with great potential in hard and soft information fusion domains, but there has not yet been adequate consideration of human interaction and hypothesis-based reasoning in CEP systems.

Multi-Agent Systems (MAS)

Software agents are commonly defined as computer systems capable of some autonomous action within a specific environment in which it is situated [I.1.16]. Agents have an extensive history, with countless publications (including [I.1.17], [I.1.18], [I.1.19], [I.1.20], [I.1.21], [I.1.22], [I.2.16], [I.1.23], [I.1.16]), numerous dedicated annual conferences (including the ACM *International Conference on Autonomous Agents and Multi agent Systems*), and multiple related IEEE and ACM journals.

Although the term "autonomous" implies a great deal of artificial intelligence capability, the autonomous actions carried out by these software agents can be relatively simple while still affording many advantages from a software engineering and logical abstraction perspective. Within the context of HHCDSM, four potential contributions of software agents and multi-agent systems (MAS) are anticipated:

1. Software agents can be used for tasking and adjudication of human vs. machine assignment to various tasks.
2. Mobile agents or "agent shells" [I.1.22] may be viable for integration of top-down or hypothesis-driven data mining within the typically bottom-up or data driven framework provided by open source CEP tools. This appears to be a novel approach and a contribution of this research.
3. Mobile agents can be used to optimize the routing of data for more efficient utilization of distributed inputs (e.g. observers, sensors), processing nodes (e.g. human analysts, machine cognition applications), data stores, and system users.
4. Agents can be used to implement Klein's Recognition Primed Decision model to support naturalistic decision-making and team cognition [I.1.19].

During this research effort, Rimland [I.5.7] developed a prototype system to utilize CEP and MAS and applied it to synthetic hard and soft data related to a complex weather monitoring task. This prototype relied on a novel architecture for integrating CEP and MAS paradigms in a manner that allows for the most useful capabilities of each approach to be applied to an analysis problem without causing undue interference or compromising the core tenets of either paradigm. The Object-to-Agent queue used to accomplish this is shown in Figure 50.

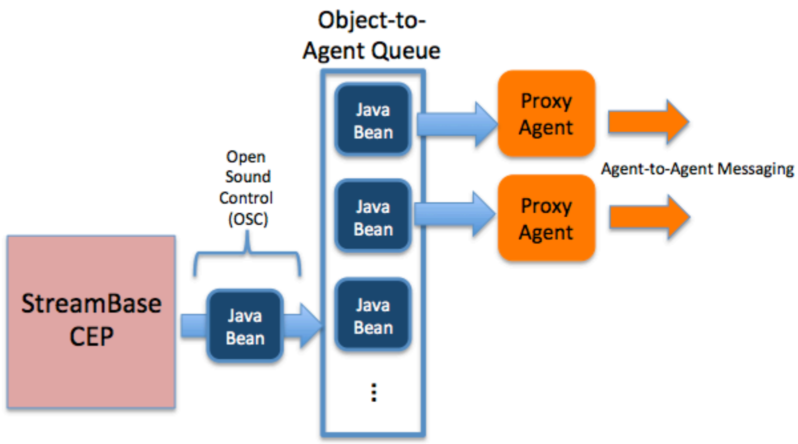


Figure 50: Object-to-Agent Queue being used as an interface between StreamBase CEP and Software Agents

Once the CEP objects have been transitioned into the multi-agent environment, the processing tasks are accomplished by the "community" of agents shown in the table below. In future prototypes, additional agents classes would be added for tasking, adjudication, human interaction, and other functionality as needed.

Table 29: A summary of software agents created for the CEP/MAS prototype

Summary of Software Agents		
Agent Name	**Java Class**	**Agent Description**
Proxy Agent	Proxy_agent.java	The proxy agents monitor the Object-to-Agent queue for the arrival of new Java Beans containing messages for the agents. It then extracts the information from that Bean and passes it on to an available Primary Receiver agent.
Primary Receiver	Recv_agent.java	The Primary Receiver agents decide which class of agent should be recruited to perform the task at hand. In this simplified prototype, the Primary Receiver always calls the Text Analysis Agent.
Text Analysis Agent	Text_analysis_agent.java	The Text Analysis agent establishes an HTTP connection to the AlchemyAPI server, sends a properly formatted HTTP POST to their web service, parses the sentiment analysis values out of the HTTP response received from their web service, and modulated the threat index based on that response.
Data Cache Agent	Data_cache_agent.java	The Data Cache agent (not used in the experiment) could optionally anticipate future queries and cache the results during periods of low system utilization.

The accompanying experiment, shown in Figure 51, was performed using this prototype with several thousand synthetic hard and soft data points.

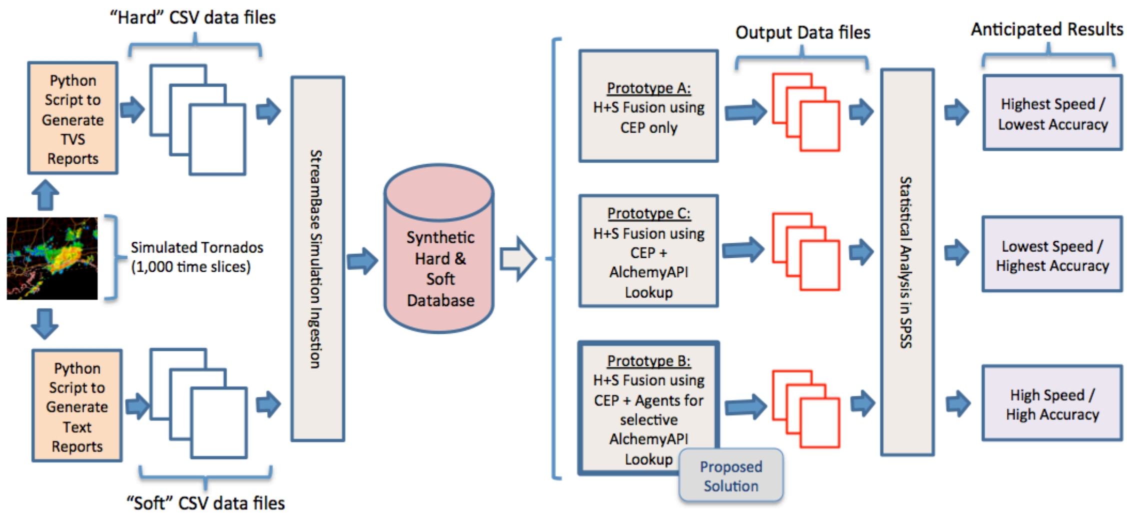


Figure 51: Overview of the experiment and summary of anticipated statistical findings

On dense datasets requiring geospatial and temporal co-registration across multiple modalities of hard and soft data, as well as consideration of message provenance, the novel CEP+MAS method performed up to 18.42 times faster than competing approaches. Complete statistical results of these experiments are available in [I.5.7].

In year five, this integration of Complex Event Processing (CEP) and Multi-Agent Systems was further extended by applying CEP to the sonification technique of vocal synthesis [I.2.1] in order to reduce large and complex heterogeneous datasets into much smaller summary datasets (see Figure 52) that were more amenable to analysis via listening.

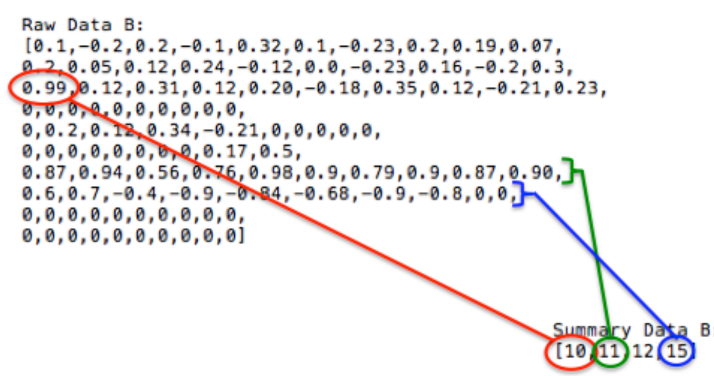


Figure 52: Reduction of dataset complexity via application of Complex Event Processing

The overall architecture for the conversion of raw data to audible is shown in Figure 53, and described in detail in [I.2.1].

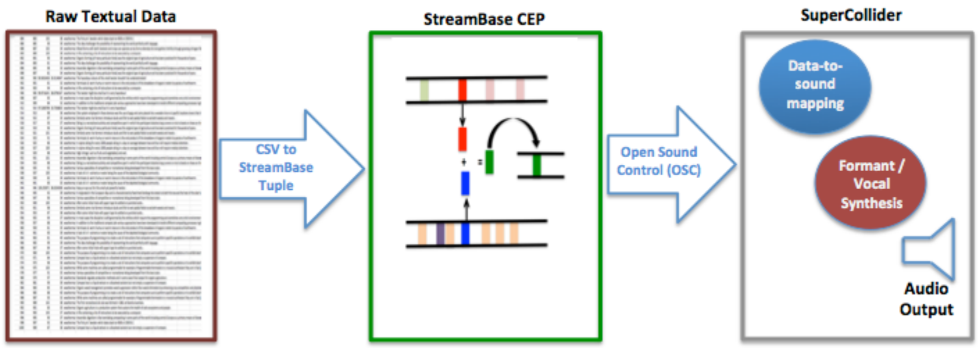


Figure 53: Raw textual data is reduced in complexity via CEP algorithms and then converted to formant/vocal sounds via SuperCollider

3.3.4 Penn State Appendix A: Summary of PSU Five Year Accomplishments

This appendix provides a summary of the Pennsylvania State University team accomplishments during this five year project. Additional details are provided in the annual reports and published papers, theses and dissertations. The appendix includes: i) a summary of key performance indicators, ii) a summary of accomplishments by year, iii) a summary of publications, and iv) a summary of outreach and technology transition activities.

Table 30: Summary of Key Performance Indicators

Summary of Key Performance Indicators						
Indicators	Year 1	Year 2	Year 3	Year 4	Year 5	Total
Student Support						
• Post-PhD					1	1
• Graduate students	3	8	6	6	5	28
• Undergraduates	4	4	4	3	1	16
Degrees Awarded						
• M.S. degrees	3	5	1	1	1	11
• Ph.D. degrees			1	1	1	3
Publications						
• Refereed conference papers	1	8	9	10	7	35
• Books and book chapters	1	2		3	4	10
• Technical reports	4	5	2	1	1	13
• Theses	3	5	2	2	2	14
Technology Transitions						
• Interactions with industry	5	4	4	8	3	24
• Interactions with govt. agencies	3	4	4	3	5	19
• SYNCOIN distribution			16	22	23	61

3.3.5 Penn State Summary of Accomplishments

The following provides a summary of annual technical accomplishments for the five year period. In general, over the five year period, the Penn State team focused on areas that included;

- *Computing and communication infrastructure* - Evaluation of current needs and the state of the art in tools, techniques and paradigms for hard and soft information fusion
- *Distributed software infrastructure* – Creation and implementation of a software infrastructure for networked integration of hard and soft fusion tools including visualization and analysis aids
- *Creation and evaluation of test data* – Analysis of existing data sets and development of a special synthetic counter insurgency (SYNCOIN) data set that includes extensive soft (message) data and hard (physical sensor data including Lidar, IR, hyperspectral, visual and related data), along with "ground truth" information on events, activities for IED scenarios.
- *Hard sensor fusion* – creation of new methods for fusing hard sensor data including 2-D and 3-D fusion at the data, feature and decision levels (e.g., target characterization, identification and tracking)
- *Analytic tools* – Design and implementation of semantic-based tools such as Complex Event Processing (CEP) and Coherence Network (CN) methods to support hypothesis generation and focus of attention for human analysts
- *Human Computer Interaction (HCI)* – Design, development and evaluation of a novel visualization toolkit for viewing temporal, geospatial and network-based relationships among entities
- *Human in the loop experiments* – Conduct of human in the loop experiments that included; i) task analysis and knowledge elicitation to understand hybrid human cognition (computer automation plus human cognition), team cognition and collaboration, ii) real-world hard and soft data collection involving humans as observers, analysts and collaborative decision-makers, and iii) evaluation of the effectiveness of visual and cognitive aids.

The results of the Penn State research have been integrated in the overall MURI team hard and soft fusion process as illustrated in Figure 54 below.

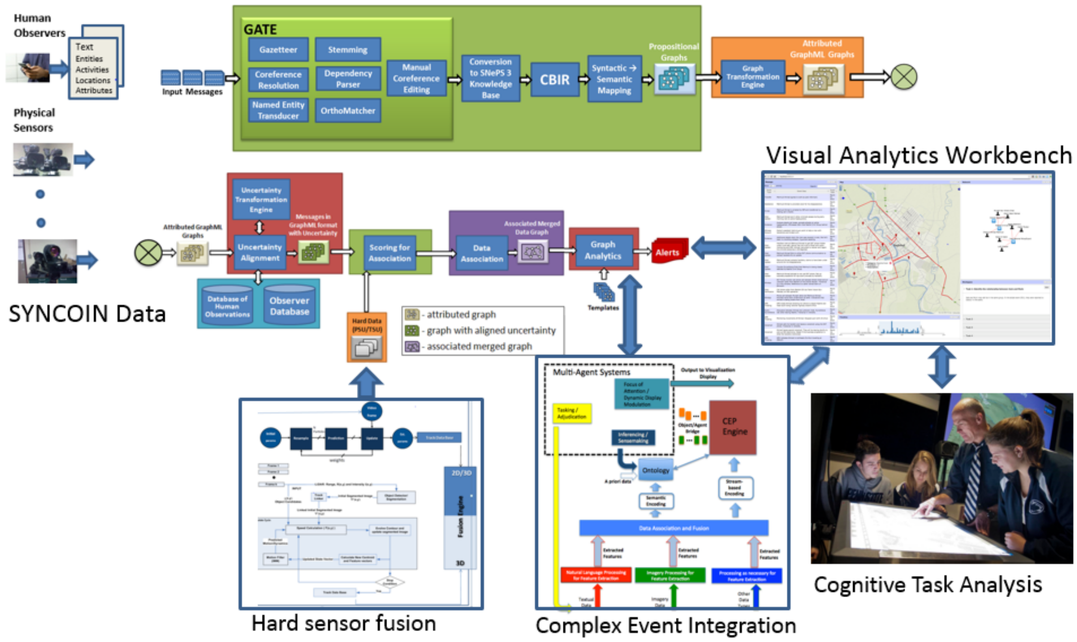


Figure 54: Penn State components of overall hard and soft fusion process

Summary of Year 1 Accomplishments

- Team formation
- Initial assessment of data needs for test and evaluation
- Evaluated 3 data sets (Hasten (DARPA), STEF (ARL), and Enhanced STEF)
- Evaluated and demonstrated initial signal and image processing algorithms for JDL Level-0 and Level-1 hard sensor fusion (e.g., Kalman Tracking, SIFT)
- Conducted initial investigation of cyber-infrastructure evolving standards (e.g., Open Geospatial Consortium Models)
- Acquired and evaluated the Fusion Exploitation Framework (FEF) from Potomac Fusion (http://www.potomacfusion.com/products/)

Summary of Year 2 Accomplishments

- Initiated development of synthetic hard/soft data set (SYNCOIN) – articulated operational perspectives on military COIN operations, developed initial soft data set (600 messages) with ground truth products
- Selected a set of hard sensors for experimentation – 3-D LIDAR, 2-D video, IR, visual and others
- Developed and demonstrated processing flow and algorithms for hard fusion processing – using MATLAB fusion/geo-mapping capability
- Implemented algorithms to fuse 3-D LIDAR and 2-D video for target ID and tracking of vehicles and humans

- Demonstrated General Dynamics GeoSuite (http://www.gdc4s.com/geosuite) mobile application for soft annotation of hard data
- Planned and conducted initial human in the loop, off-campus-based experiments at Pleasant Gap Fire Safety Training Facility (https://www.facebook.com/pages/Centre-County-Public-Safety-Training-Center/172746352759034)
- Implemented an initial infrastructure for integration/transition
- Began dissemination of SYNCOIN data to the data fusion community.

Summary of Year 3 Accomplishments

Fusion of hard sensor data – Implemented algorithms for fusion of hard sensor data

- Developed prototype applications for target identification, localization and tracking in MATLAB and C++
- Implemented MATLAB fusion/geo-mapping capability
- Explored Situation Awareness Dashboard application using the Command Post of the Future (CPOF)
- Test and evaluation – Developed a test and evaluation (T&E) approach progressing from synthetic hard and soft data set to human experiments
    - Conducted 3, multi-day data collection events at the Centre County Public Safety Training Center in Pleasant Gap, PA (https://www.facebook.com/pages/Centre-County-Public-Safety-Training-Center/172746352759034) using COIN mini-vignettes involving multiple hard sensors and human actors
    - Refined the development of SYNCOIN, a synthetic hard and soft data set including interlaced scenarios, 600 text messages and synthetic hard data including; mapped PIRs to I&W to SYNCOIN messages and Linked physical data (from Pleasant Gap collects) to SYNCOIN threads
    - Created ground truth products (utilizing Analyst Notebook) to check the veracity of fusion processes

Integration and transition – Designed and implemented an integration & transition environment.

- Developed baseline information architecture and service oriented architecture approach for integration, test and transition,
- Implemented and demonstrated proof-of-concept service oriented architecture (SOA)
- Acquired, assessed and implemented the Fusion Exploitation Framework (FEF) transition environment at Penn State I
- Developed proof of concept system to encode/decode/transmit hard/soft data in OGC-compliant formats
- Investigated technologies, standards, and applications
- Continued dissemination of SYNCOIN data to the data fusion community

Summary of Year 4 Accomplishments

Synthetic hard/soft data set

- Continued evolution of the SYNCOIN data set including; a new physical sensor data and new soft and hard sensor data links and meta-data
- Conducted human in the loop cognitive task analyses
- Continued dissemination of SYNCOIN data to the data fusion community

Hard sensor data fusion

- Continued development of new algorithms for fusion of hard sensor data
- Range imaging tracking, (Interacting Multiple Mode (IMM) Kalman Filters) Tracking
- VNIR Color Particle Filter Tracking – VNIR Image fusion and Multi-Model Object characterization including Range/Depth Automated Segmentation Algorithm

New automated inference tools

- Created intelligent agents for improved focus of analyst attention
- Applied Complex Event Processing (CEP) to SYNCOIN data
- Developed a novel technique for integrating CEP) with Multi-Agent Systems(MAS)

Visualization toolkit

- Implemented web-based interactive visual analysis (IVA) toolkit
- Created relational database to link SYNCOIN geo, temporal and human network data

Integration & network based processing

- Robust cyber-infrastructure for distributed H/S processing - StreamBase CEP, Advanced message queuing protocol (AMQP), RabbitMQ, Open Geospatial Consortium standards for TML and Event Pattern Markup Language, AchemyAPL, RDF/OWL
    - Demonstrated SYNCOIN data in AXISPro (http://www.textronsystems.com/products/advanced-information/axis-pro)

Summary of Year 5 Accomplishments

Test and Evaluation

- Continued refinement of the SYNCOIN data fusion test set including refinement of the soft data set to meet emerging evaluation requirements and enhancement of the hard sensor meta-data to assist correlation and association with graph matching demonstrations and evaluation.
- Participation in the July 2014 meeting of the International Society of Information Fusion (ISIF) Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG)

Cyber Infrastructure

- Continued evolution and evaluation of cyber infrastructure techniques and tools to facilitate distributed hard and soft information fusion
- Extended the [I.1.24] framework for distributed context-aware sensing, fusion and decision-making applications in a cloud environment.

Hard sensor data fusion

- Completed automated classification of human forms from 2-D and 3-D depth map fused data
- Designed and implemented a multi-level association (translation of parametric data to "story-book" scene characterization)
- Final data collect using KINECT – merged color tracker and depth map tracker

Automated sense-making algorithms

- Applied and extended Complex Event Processing/Multi-Agent Systems to SYNCOIN data
- Developed a coherence network algorithm for SYNCOIN
- Linked Complex Event Processing/Multi-Agent Systems and Coherence Network processing

Visual analytics and cognitive assessment

- Extended the cognitive task analysis to improve understanding of analytical reasoning processes
- Extended the current toolkit (e.g., text analysis, recommender support, etc.) for enhanced analysis capabilities
- Investigated techniques to enable collaborative analysis with multiple distributed analysts

- Integrated visual interaction (human-supported cognition) with computer automated processing
- Implemented multi-view data exploratory analysis (data dashboard) with several view coordination strategies
- Developed a novel sonification technique using CEP and vocal synthesis
- Integrated visual data fusion with intelligence analysis (Dashboard + workbench)
- Fielded a beta version of the toolkit and initiated conduct human in the loop experiments\

3.3.6 Summary of Publications

The following is a list of publications related to the five year MURI project. The publications are listed in the following categories; refereed conference proceedings, books and book chapters, technical reports, theses and dissertations, presentations and organized conference sessions.

Refereed Conference Proceedings

- Sherry, R., J. Gabor and D. Hall (2009), "Information fusion to support real time accident diagnosis and accident management", *Proceedings of the OECD/NEA Workshop on Implementation of Severe Accident Management (SAM) Measures*, October 26-28, 2009, Schloss Bottstein, Switzerland

- Hall, D., Hellar, B., and McNeese, M. D., (2009), "The Extreme Events Laboratory: A cyber infrastructure for performing experiments to quantify the effectiveness of human-centered information fusion," *Proceedings of the 2009 International Conference on Information Fusion (Fusion 2009),* Seattle, Washington, July, 2009
- D. Hall, J. Graham, L. More and J. Rimland (2010), "Test and evaluation of soft/hard information fusion systems: an experimental environment, methodology and initial data sets", *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, UK, July, 2010

- N. Giacobe, (2010), "Mining social media in extreme events: lessons learned from the DARPA network challenge", *Proceedings of the IEEE Conference on Homeland Securities Technologies (IEEE HST 2010),* Waltham, MA, November 2010

- J. Llinas, R. Nagi, D. Hall and J. Lavery (2010), "A multidisciplinary university research initiative in hard and soft information fusion: overview, research strategies and initial results," in *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, UK, July, 2010

- R. Tutwiler, D. J. Natale, M. S. Baran, R. L. Tutwiler, (2010), "Live motion 3D data processing", *Proceedings of the IDGA 9th Image Fusion Summit,* November 15 - 17, 2010, Sheraton Premiere at Tysons Corner, Vienna, VA

- J. Graham, J. Rimland, D. Hall, (2011), "A COIN-inspired synthetic data set for quantitative evaluation of hard and soft fusion systems", *Proceedings of Fusion 2011: the International Conference on Information Fusion,* Chicago, IL, July, 2011
- R. Tutwiler, M. Baran, D. Natale, C. Griffin, J. Daughtry, M. McQuillan, J. Rimland, and D. Hall, (2011), "Hard sensor fusion for COIN inspired situation awareness", *Proceedings of Fusion 2011: the International Conference on Information Fusion,* Chicago, IL, July, 2011
- J. Rimland, (2011), "A multi-agent infrastructure for hard and soft information fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011
- J. Rimland, (2011), "JDL level 0 and 1 algorithms for processing and fusion of hard sensor data", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011
- J. Graham (2011), "A new synthetic dataset for evaluating soft and hard fusion algorithms", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011
- D. J. Natale, M. S. Baran, R. Tutwiler and D. L. Hall, (2011), "3DSF: three dimensional spatio-temporal fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011
- D. Hall, (2011) invited panel discussion, "Real world issues and challenges in hard and soft data fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25 April, 2011
- J. Graham, J. Rimland and D. Hall (2011), "A COIN-inspired synthetic data set for qualitative evaluation of hard and soft fusion systems", *Proceedings of the 14th International Conference on Information Fusion*, Chicago, IL, July, 2011
- D. Hall, G. Iyer, M. Ballora, R. Cole, H. Kruesi and H. Greene, (2011), "Use of auditory displays in anomaly detection, *Proceedings of the National Symposium on Sensor and Data Fusion*, Oct. 24-28, 2011
- M. S. Baran, C. J. Natale, R. Tutwiler, M. McQuillan, C. Griffin, J. Daughtry, J. Rimland and D. Hall (2011), "Hard sensor fusion for COIN inspired situation awareness", *Proceedings of the 14th International Conference on Information Fusion*, Chicago, IL, July, 2011
- D. Hall, G. Iyer, M. Ballora, R. Cole, H. Kruesi and H. Greene, (2011), "Use of auditory displays in anomaly detection, *Proceedings of the National Symposium on Sensor and Data Fusion*, Oct. 24-28, 2011

- J. Rimland, D. Hall and J. Graham, (2012), "Human cognitive and perceptual factors on JDL level-4 hard/soft fusion", *Proceedings of the SPIE Conference on Multi-sensor, Multisource Information Fusion: Architectures, Algorithms*, and Applications 2012, Baltimore, MD, April 23-27, 2012
- D. Sudit, S. Kumara and D. Hall, (2012), "Complex mathematical model for soft processes in information fusion," *Proceedings of the ISERC 2012 Conference*, Orlando, FL, April, 2012
- J. Graham and D. Hall, (2012), "The use of Analytic Decision Game (ADG) methods for test and evaluation of hard and soft data fusion systems and education of a new generation of data fusion analysts," *Proceedings of the National Symposium on Sensor Data Fusion (NSSDF),* Washington, DC, 22-16 October, 2012
- D. Kretz, B. Simpson and J. Graham, (2012), "A Game-Based Experimental Protocol for Identifying and Overcoming Judgment Biases in Forensic Decision Analysis", *IEEE International Conference on Technologies for Homeland Security*, Waltham, MA, 13-15 November, 2012.
- Matthew S. Baran; Richard L. Tutwiler; David L. Hall; Donald J. Natale "Surface reconstruction for 3D remote sensing", *Proceedings of SPIE 2012*, Baltimore, MD
- J. Rimland, D. Coughlin, D. Hall and J. Graham, (2012),"Advances in data representation for hard/soft information fusion", *Proceedings of SPIE 2012*, Baltimore, MD
- J. Rimland and J. Llinas, (2012), "Network and infrastructure considerations for hard and soft information fusion processes,", *Proceedings of the International Society of Information Fusion*, FUSION 2012, July, 2012, Singapore
- J. Rimland, M. Ballora and D. Hall, (2013), "Hard and soft information fusion in sonification for assistive mobile device technology", *Proceedings of the International Conference on Auditory Display* (ICAD_2013), July 6 – 10, 2013, Lodz University of Technology, Poland
- M.S. Baran, R.L. Tutwiler, D. J. Natale, M .S. Bassett, M. P. Haner, (2013), "Multi-Modal Detection of Man-Made Objects in Simulated Aerial Imagery", *Proc. SPIE*. 8743*, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultra-spectral Imagery XIX* 87430P (May 18, 2013)
- N. S. Butler, R.L. Tutwiler, (2013), "Invariant unsupervised segmentation of dismounts in depth images", *Proc. SPIE*. 8745, *Signal Processing, Sensor Fusion, and Target Recognition* XXII 87451B (May 23, 2013)
- J. Rimland and M. Ballora ,(2013), "Beyond visualization of Big Data: a multi-stage data exploration approach using visualization, sonification, and storification", *Proceedings of SPIE 2013*.
- J. Rimland, M. McNeese and D. Hall, (2013), "Conserving Analyst Attention Units: Use of Multi-agent Software and CEP Methods to Assist Information Analysis", *Proceedings*

of SPIE DSS Conference on Next-Generation Analyst, vol. 8758, April, 2013, Baltimore, Md.

- J. Rimland and M. Ballora, (2014), “Using Complex Event Processing (CEP) and vocal synthesis techniques to improve comprehension of sonified human-centric data,” *Proceedings of SPIE 2014,* Baltimore, MD, May 6, 2014.
- J. Rimland and M. Ballora, (2014), “Using vocal-based sounds to represent sentiment in complex event processing”, *Proceedings of the International Conference on Auditory Display* (ICAD), 2014
- J. Rimland, S. Shaffer and D. L. Hall, (2014) "A hitchhiker’s guide to distributed hard and soft information fusion infrastructure development", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain.
- S. Shaffer, (2014), “Automatic theory generation from analyst text files using coherence networks”, *Proceedings of the SPIE 2014 Conference*, Baltimore, MD, May 6, 2014.
- G. Cai and J. Graham (2014), “Semantic data fusion through visually-enabled analytical reasoning”, *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain
- G. Cai, G. Gross, J. Llinas and D. Hall (2014), “A visual analytic framework for data fusion in investigative intelligence”, *Proceedings of the SPIE volume 9122, Next-Generation Analyst II.* Baltimore, Maryland, USA
- John P. Morgan, Richard L. Tutwiler, “ Real-Time reconstruction of depth sequences using signed distance functions ”, *Proceedings of the SPIE 2014 Conference*, Baltimore, MD, May 6, 2014

Books and Book Chapters

- Hall, D. and Jordan, J. (2010). *Human Centered Information Fusion*. Artech House, Inc.
- D. Hall and S. Aungst (2010), The use of soft sensors and I-Space for improved combat ID, chapter 10 in *Human Factors in Combat Identification*, ed. by D. Andrews, R. Herz and M. Wolf, Ashgate, pp 161-170
- D. Hall, J. Llinas, C. Chong, K. C. Chang, editors, (2012), *Distributed Data Fusion for Network-Centric Operations,* CRC Press
- D. Hall, (2012), “Understanding the new users: collaborative decision-making paradigms, communities of interest, and complex adaptive systems”, chapter 3 in D. Hall, J. Llinas, C. Chong, K. C. Chang, editors, *Distributed Data Fusion for Network-Centric Operations,* CRC Press, 2012

- D. L. Hall, C. M. Hall, S. A. H. McMullen and M. McMullen (2011), "Improving uncertainty assessment and situational awareness using hard and soft information fusion", in *Risk Management in Decision Making: Intelligent Methodologies and Applications,* edited by J. Lu, Springer, 2011
- D. Hall, (2011), "The Emergence of Human-Centric Information Fusion," chapter 18 in *Distributed Sensor Networks*, 2nd edition, 2011
- D. L. Hall, 2012), "Perspectives on Distributed Data Fusion", chapter 1 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press, August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang
- J. Rimland, (2012), "Service-Oriented Architecture for Human-Centric Information Fusion," chapter 13 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press, August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang
- D. Hall, (2012), "The Emergence of Human-Centric Information Fusion," chapter 27 in Distributed Sensor Networks, 2nd edition, 2012, edited by S. Iyengar and R. Brooks
- D. L. Hall, (2012), "Perspectives on Distributed Data Fusion", chapter 1 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press, August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang

Technical Reports

- D. Saab and F. Fonseca, (2009), *Participatory Sensing: A Review of the Literature and State of the Art Practices,* Technical Report for the Penn State University Center for Network Centric Cognition and Information Fusion (NC2IF), November 11, 2009 (78 pages)
- D. L. Hall and M. McNeese (2010), *First year interim report for the Multidisciplinary University Research Initiative (MURI) on Unified Research on Network-based Hard/Soft Information Fusion*, prepared for the U. S. Army Research Office, August, 23, 2010
- R. L. Tutwiler, *MURI Hard Sensor Fusion Performance Characterization*, Technical report, May, 2011
- J. Graham, *SYNCOIN Data Set*, Technical report, December, 2010
- J. Rimland, (2011) "Factors determining success in participatory sensing campaigns.", Internal Report for the NC^{2}IF Research Center, January, 2011
- J. Rimland, (2011), "Cognitive factors in data fusion and visualization", Internal Report for the NC^{2}IF Research Center, March, 2011

- J. Rimland, (2011), “The role of perceptual factors in human-in-the-loop HCI”, Internal Report for the NC^{2}IF Research Center, May, 2011
- D. Hall, J. Graham, M. McNeese, J. Rimland and R. Tutwiler (2011), Second Year Interim Progress Report: *Army Research Office Multidisciplinary University Research Initiative (MURI) grant on Unified Research on Network-based Hard/Soft information Fusion*, August 23, 2011
- D. L. Hall, J. Graham, M. McNeese, J. Rimland and R. Tutwiler, (2012) , *Third Year Interim Progress Report: Army Research Office Multidisciplinary University Research Initiative (MURI) grant on Unified Research on Network-based Hard/Soft information Fusion*, August 23, 2012 (28 pages)
- R. L. Tutwiler, *MURI Hard Sensor Fusion Performance Characterization*, Technical report, May, 2011
- J. Graham, *SYNCOIN Data Set*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC^{2}IF) Research Center, IST Building, University Park, PA 16802, revised, December, 2011
- J. Graham, Scene *Setter for MURI Demonstration*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, July 30, 2012
- N. Giacobe, *SYNCOIN Word Clouds*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802 May 1, 2012
- D. L. Hall, J. Graham, M. McNeese, J. Rimland, R. Tutwiler and G. Cai, (2013), *Unified Research on Network-based Hard/Soft Information Fusion*, Interim progress report for the Army Research Office Multidisciplinary University Research Initiative, July, 2013, (42 pages)
- J. Graham et al, (2014), *Analyst Workbench Instructional Guide,* Technical report for the NC^{2}IF Research Center, August, 2014 (30 pages)

Theses and Dissertations

- K. Misra, (2010) A cyber infrastructure for hard and soft data fusion, M. S. thesis in Electrical Engineering, The Pennsylvania State University, University Park, PA
- Xu M.S. (2010) Unsupervised flow-level clustering in network anomaly detection, M. S. thesis in Electrical Engineering, The Pennsylvania State University, University Park, PA
- Rachana Reddy Agumamidi, (2011) ,M. S. thesis, The Pennsylvania State University, Electrical Engineering, “Hard Sensor Processing for Data Fusion”, May, 2011
- Ganesh Iyer, (2011), M.S. thesis, The Pennsylvania State University, Electrical Engineering, “Approaches to hard and soft sensors’ data fusion”, June, 2011

- A. Godbole, (2013) Improving utilization of mobile device technology for distributed emergency teams, M.S. theses in Computer Science and Engineering, The Pennsylvania State University, June 2013
- J. C. Rimland (2013), Hybrid human-computing distributed sense-making: Extending the SOA paradigm for dynamic adjudication and optimization of human and computer roles", Ph.D. dissertation in Information Sciences and Technology, The Pennsylvania State University, August, 2013

Presentations

- D. Hall, (2010), Human Centered Fusion: The Emerging Role of Humans in Situation Awareness, Keynote presentation at SPIE Conference on Defense, Security and Sensing, April 5-9, 2010, Orlando
- D. L. Hall , "Asymmetric Information Warfare: Challenges and Opportunities in Information Fusion," keynote presentation at the 2012 DoDIIS Worldwide Conference, April 2nd, 2012, Denver, CO
- D. L. Hall (2011), invited participation in ETUR Panel: "Developments and issues in uncertainty representation", FUSION 2011: International Society of Information Fusion, Chicago, Ill, July 6, 2011
- J. Graham, "SYNCOIN: a synthetic dataset for evaluating hard and soft fusion algorithms," presentation to SI Org University Innovation Day Share [IT], 2 August 2012, Chantilly, VA

Organized conference sessions

- Next Generation Analyst: Special one-day session organized for the SPIE Conference on Sensing Technology and Applications, May 2012 in Baltimore, MD
- Next Generation Analyst II: Special one-day session organized for the SPIE Conference on Sensing Technology and Applications, May 2013 in Baltimore, MD
- Next Generation Analyst III: Special one-day session organized for the SPIE Conference on Sensing Technology and Applications, May 2014 in Baltimore, MD (proposed)
- International Society of Information Fusion - FUSION 2014 Conference held in Salamanca, Spain, July, 2014 – organized a special session on advances in hard and soft information fusion

3.3.7 Summary of Outreach and Technology Transition Activities

Throughout the MURI project, the Penn State team has been active in seeking collaboration and opportunities for technology transition to both industry and to government agencies. Table 31 provides a summary of those interactions over the course of the five year program.

Table 31: Summary of Industrial and Government Interactions

Year	Interactions with Industry	Interaction with Govt. Agencies
1	• Collaborations with Penn State Police Services regarding stadium protection • General Dynamics (exploring potential Command Post of Future collaboration) • Lockheed Martin (planning related IR&D project) • i2 corporation (interaction with Analyst Notebook) • Mechdyne discussions for possible collaboration for advanced 3-D visualization	• Discussion with NAVSEA Warfare Centers (NSWC Crane) • Proposed Red Cell collaboration effort with Kira Hutchinson (from JIEDDO) • Centre County Emergency Management Services
2	• Collaborations with Penn State Police Services regarding stadium protection • General Dynamics C4 Systems (obtained copy of the Command Post of Future and Tactical Ground Reporting System (TIGR) and collaborated on use of GeoSuite) • i2 corporation (interaction with Analyst Notebook) • Mechdyne discussions for collaboration for advanced 3-D visualization	• Discussions with NAVSEA Warfare Centers (NSWC Crane) • Proposed Red Cell collaboration effort with Kira Hutchinson (from JIEDDO) • Centre County Emergency Management Services • Discussions with USAF NORTHCOM regarding Homeland Security
3	• Collaborations with Penn State Police Services regarding local major events • General Dynamics C4 Systems (exploring potential Command Post of Future and Tactical Ground Reporting System (TIGR) collaboration) • Lockheed Martin (planning related IR&D project) • i2 corporation (interaction with Analyst Notebook)	• Discussion with NAVSEA Warfare Centers (NSWC Crane) • Proposed Red Cell collaboration effort with Kira Hutchinson (from JIEDDO) • Centre County Emergency Management Services • Discussions with USAF NORTHCOM regarding Homeland Security
4	• Distributed SYNCOIN to 11 organizations and individuals • Raytheon Corporation • QuinetiQ (UK) • Aptima • VIStology Inc. • Saffron Technology • Arctan • Overwatch	• Network Science Collaborative Technology Alliance (U. of Illinois) • Institute for Defense Analysis • Evaluation of Techniques for Uncertainty Representation Working Group (ISIF)
5	• Raytheon Corporation • Boeing Corporation • Penn State Applied Research Laboratory	• MIT Lincoln Laboratory • Naval Surface Warfare Center (Crane Division) • Department of Homeland Security (DHS) • DHS/Transportation Security Administration • Washington Homeland Security Roundtable

In addition to these interactions, extensive activity has involved sharing the SYNCOIN data and associated documentation with industrial and government partners. Below is a list of the contacts with whom we have shared the SYNCOIN data set.

- Peter Willet, University of Connecticut
- Gavin Powell, ADS Innovation Works, UK, government technical area lead for TA 6 - Distributed Coalition Information Processing for Decision-Making
- David Nicholson, BAE Systems, London, UK
- David Dearing, Stottler Henke Associates
- David Braines, Hursley Emerging Technology Services
- Erick Blasch, Air Force Research Laboratory Sensors Directorate (AFRL/SNAA)
- Marco Pravia, BAE Systems
- Kamal Premaratne, University of Miami
- James Law, SPAWARSYSCEN – U. S. Navy Space and Naval Warfare Systems Center
- Chase Cotton, Network Science Collaborative Technology Alliance Program (CTA), U. S. Army Research Laboratory
- ETURWG – Evaluation of Techniques for Uncertainty Representation Working Group, International Society of Information Fusion (ISIF)
- International Technology Alliance
- Brian Simpson, Raytheon Corporation
- Simon Maskell, QinetiQ, UK
- Charlotte Shabarkh, Aptima, Woburn, MA
- Brian Ulicny, VIStology, INC, Framingham, MA
- Dr. Joan Carter, Institute for Defense Analysis, Alexandria, VA
- Network Science Collaborative Technology Alliance, University of Illinois, Champaign, IL
- Jim Fleming, Saffron Technology, Cary, NC
- Charles Moorefield, Arctan, Arlington, VA
- Rick Beckett, Overwatch, Textron, Philadelphia, PA
- Dr. Tony Penza, MIT Lincoln Laboratories
- Naval Surface Warfare Center, Crane Division

3.3.8 Penn State Appendix B: Additional Details on Hard Sensor Processing[10]

Processing of the collected data continued with the development and refinement of algorithms for target tracking and characterization/identification.

Segmentation and Classification from Depth Maps

Our goal is to reliably detect and segment people from raw lidar data and then feed this segmentation to a tracker that has already been developed. Here we focus on the segmentation.

In order to detect people in the depth map we need a set of invariant features that succinctly describes the depth at any location. These features can then be used with any training algorithm, such as a neural network, naïve Bayes classifier or support vector machine (SVM). The features we currently use for classification are the depth values of an arbitrarily sized 20x40 window centered on the point of interest. The initial training set is selected by hand; some samples are shown in Figure 55.

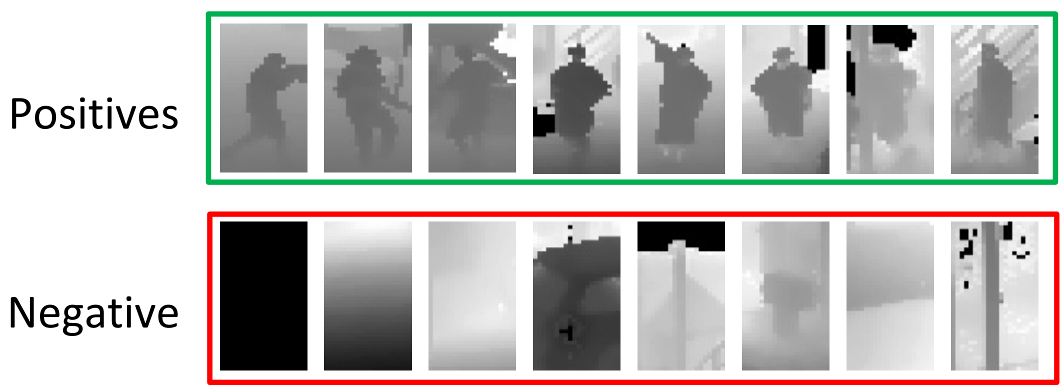


Figure 55: Some hand selected training examples

Using a 20x40 image yields an 800 element feature vector; however using a feature such as this poses a few problems. The first is scale, which we take care of by setting each window to 3x6ft (about the size of a person). The second is orientation. To robustly train and classify these images we need the person to be at about the same position and orientation in each window. We plan this by generating a structural tensor whose semi-axes can be used to orient the window at each point of interest.

The entire process is shown in Figure 56 and with the final goal in mind we now describe each step of the process.

Step 1: Read in next lidar frame

[10] Section 3.3.8 contains extensive information provided in the 4th year report, but is included here for understanding the hard sensor processing flow and algorithms

Each lidar frame consists of two 128x128 images, the intensity and depth. The intensity represents the strength of the received signal at each point. This varies from frame to frame so we only use it to threshold the depth initially by getting rid of values with low intensity. To further clean up the raw depth we smooth it with a bilateral filter and remove sparse points, without this step the gradients become messy and the structural tensors do not fit the data well. The raw and filtered depth maps are shown in Figure 57.

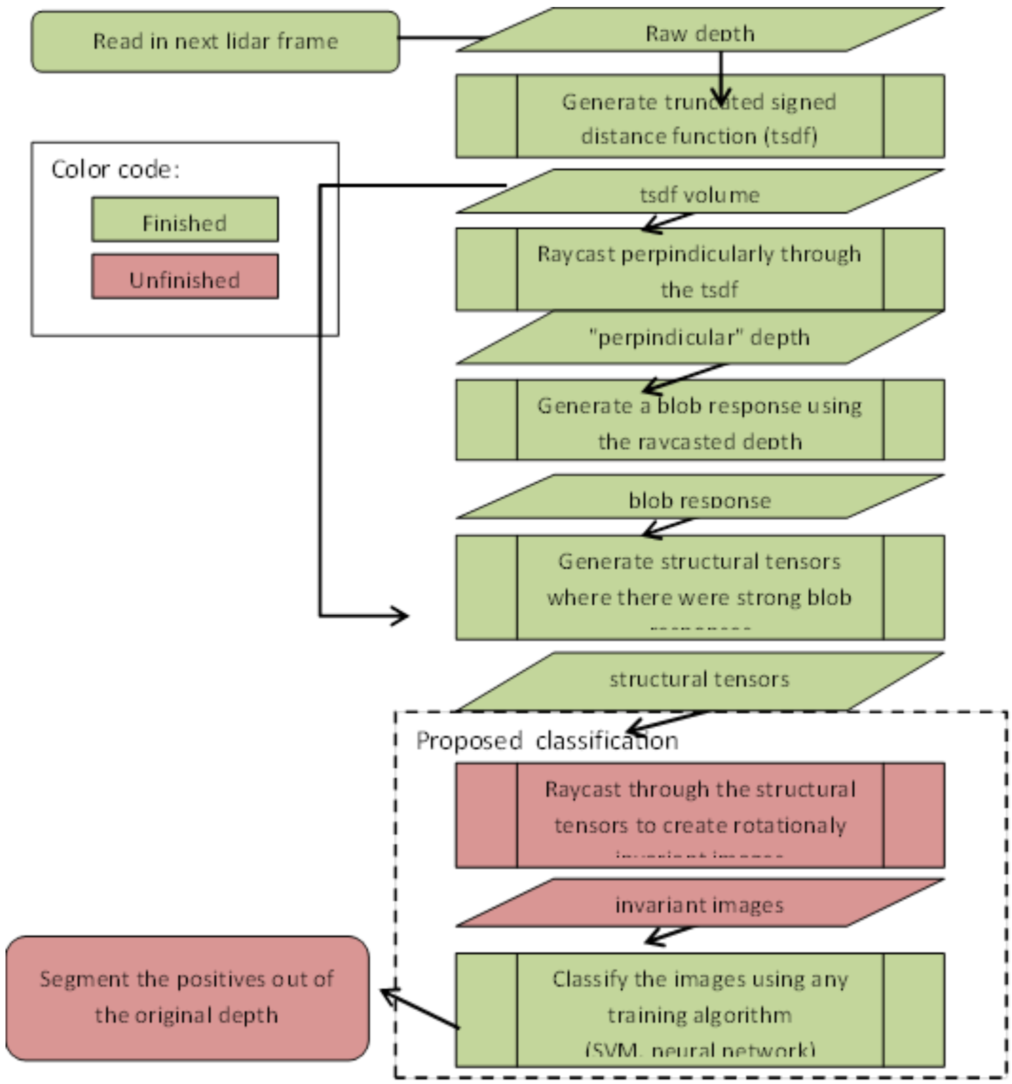


Figure 56: Current segmentation pipeline

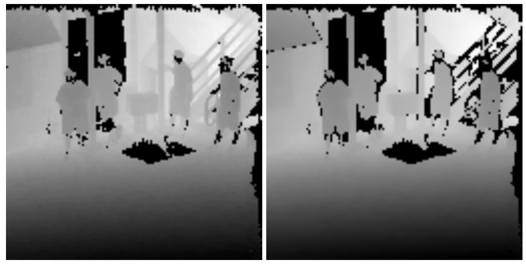
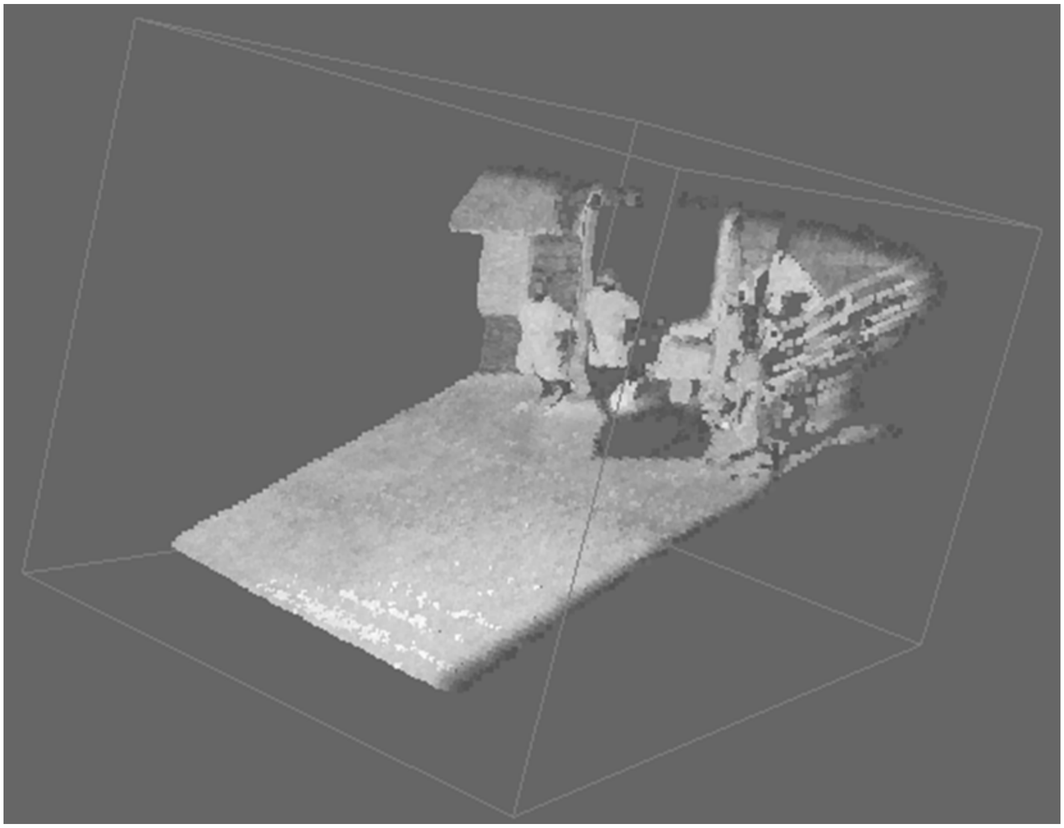


Figure 57: Raw depth A before and B after pre-processing

Step 2: Generate the truncated signed distance function

From here the depth is used to create a truncated signed distance function (TSDF) volume. We pre-define a 3D grid of voxels and then estimate the distance to the nearest surface at each voxel. The values are estimates because we only calculate distances along the ray from the viewpoint to save time. Any distance greater than a user defined threshold gets truncated to a 1, representing empty space. This threshold is arbitrary though if too large the distances can become inaccurate and if too small (with respect to the voxel density) we may not get a zero crossing. We chose to use a voxel size of .04x.04x.04m and a truncation distance of .5m. Distances in front of the surface are positive whereas behind it they are negative; this is portrayed in Figure 58 B as blue-green for positive distances and yellow-red for negative.

The TSDF volume exhibits two qualities that we take advantage of. First, by using tri-linear interpolation we can raycast through the volume to create depth maps of any size. Raycasting a depth map is done by generating a ray from the viewpoint to an arbitrary x, y, z point. We then check to see if the ray intersects the volume; if it does we find the distance along the ray from the viewpoint where it enters and exits the volume. Starting from the entry point we step along the ray with some arbitrarily small step size, we use one tenth the voxel size, until we hit a surface or exit the volume. The surface exists where the distance is zero; the point where the ray intersects a surface is called the zero crossing. Because the distance is positive in front of the surface and negative behind it we can detect the zero crossing as the point where sign changes along the ray. To create a depth map the z value of the zero crossing is recorded at the coordinates of the x, y, z point in the image plane (see Figure 59).



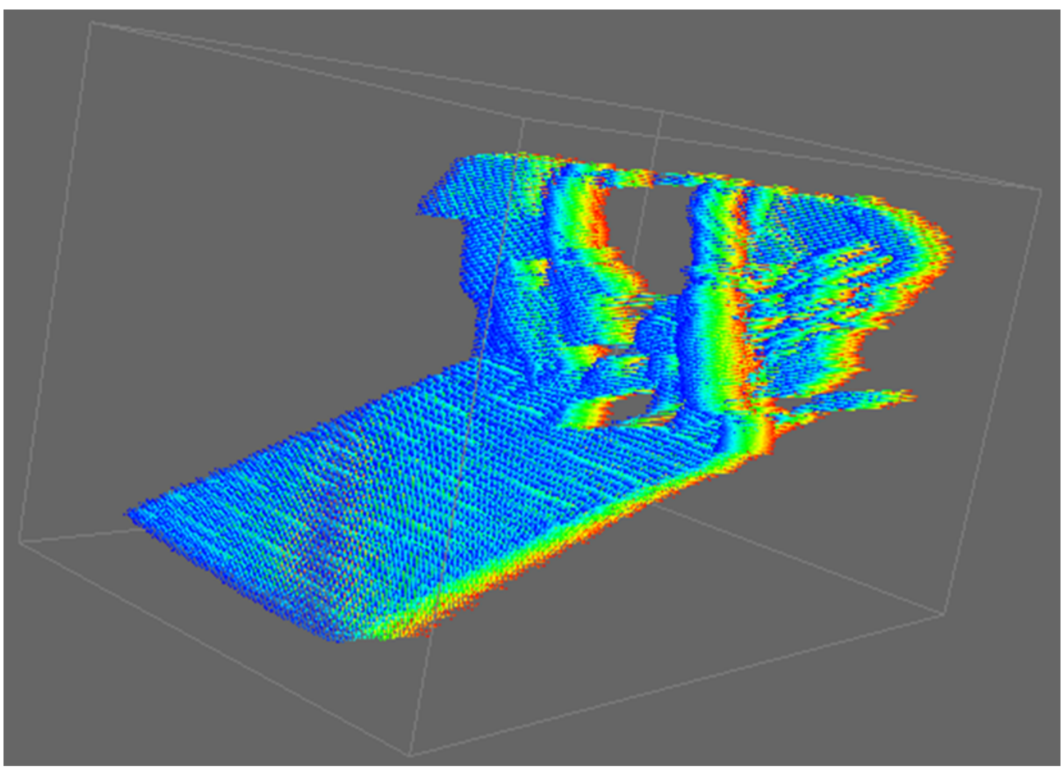


Figure 58: A- the point cloud and B- its corresponding TSDF volume

Second, the TSDF makes calculating the gradient at any point simple by using the following formula:

$$gradient_x = tsdf(x - \Delta x, y, z) - tsdf(x + \Delta x, y, z)$$
$$gradient_y = tsdf(x, y - \Delta y, z) - tsdf(x, y + \Delta y, z)$$
$$gradient_z = tsdf(x, y, z - \Delta z) - tsdf(x, y, z + \Delta z) \quad (1)$$

This is very useful when generating structural tensors later on.

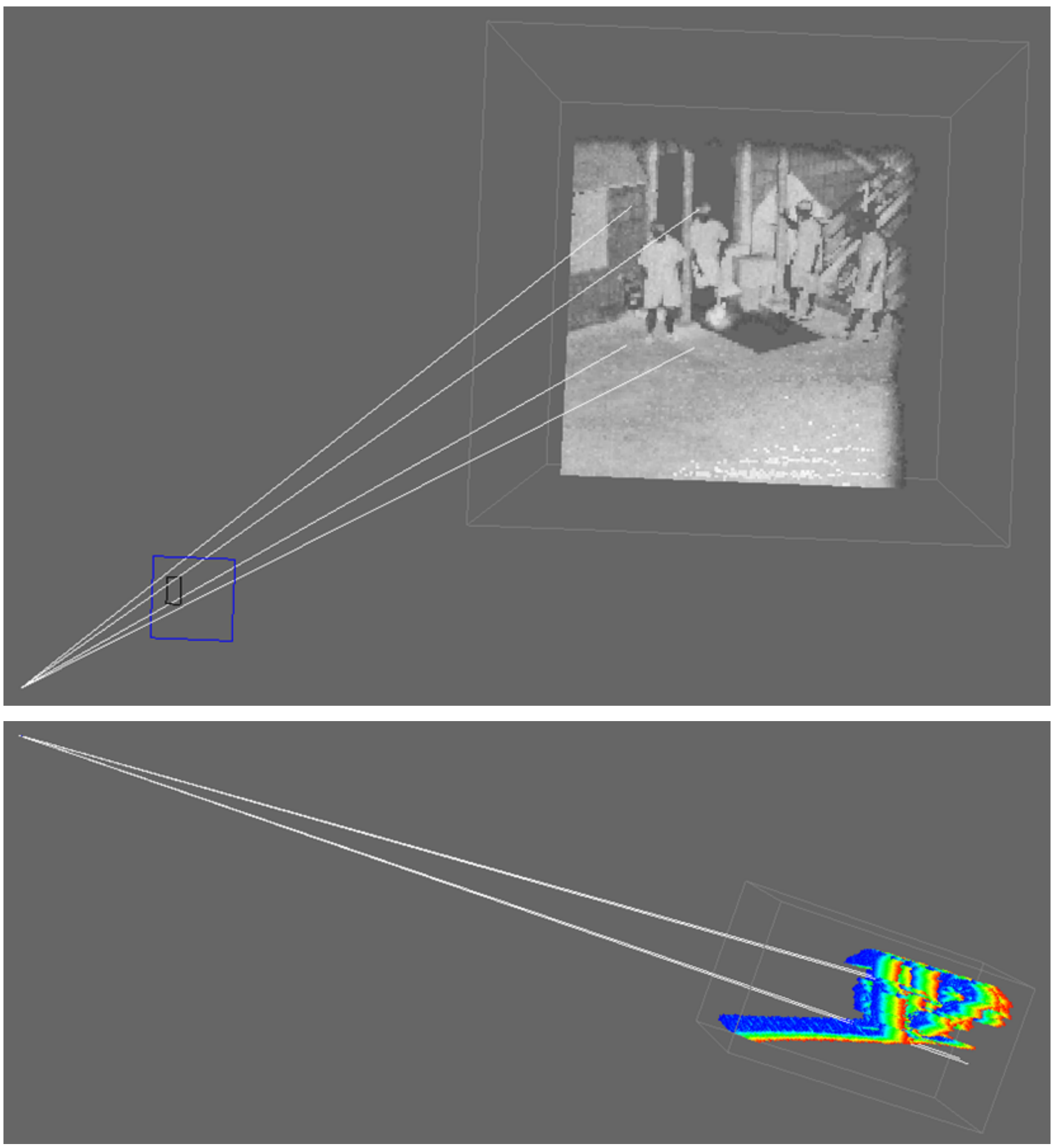
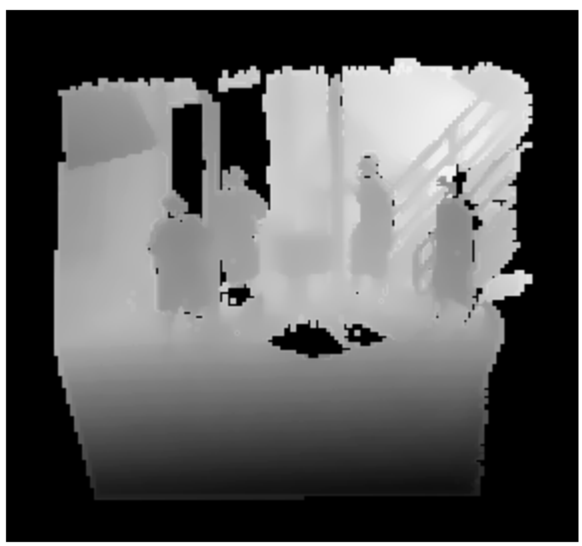


Figure 59: A- the 128x128 lidar image plane is shown in blue with the raycasted region within the image shown in black. B- the same raycast shown going through the TSDF volume. C- the resulting 20x40 depth map

Step 3: Raycast perpendicularly through the TSDF volume

The depth gives us 3D data which necessitates 3D convolution kernels for the next few steps. In our case these convolutions aren't separable and would take a long time to do at every point. However, we can take advantage of the fact that a depth map is actually "2.5D," meaning that there are no points behind other points when looking from the viewpoint. By raycasting perpendicularly (along the z-axis) through the TSDF volume we can generate a depth map where the x and y coordinates in the image hold physical meaning. Because the data is 2.5D very little information is lost in the process.



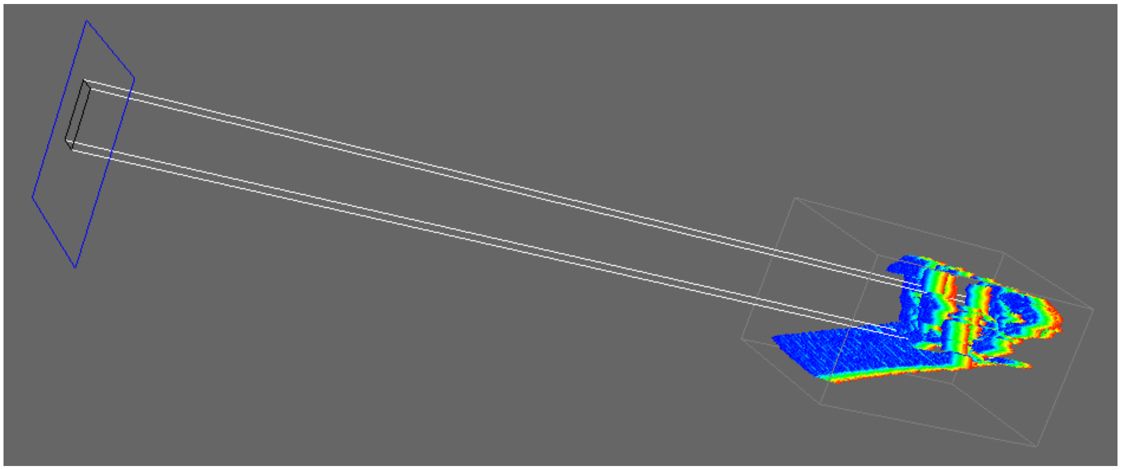


Figure 60: A- perpendicularly raycasted depth using the TSDF volume in Figure 57. B- the example in Figure 58 shown as a perpendicular raycast

Using the perpendicular depth we can do a 2D convolution in place of the 3D convolution by realizing that there is only one z value for each x, y coordinate, all of which can be interpreted as an actual distance. For example, Figure 60 shows a perpendicular depth map where each pixel is 2x2cm, which is determined by the distance between each ray when we raycast.

Step 4: Generate a blob response using the perpendicular depth

The depth map contains 42,785 non-zero depths. Attempting to generate a structural tensor at and classify every single point would take too long so we use a blob response to pick out a handful of regions that are about the size of a person. The blob response is generated by convolving the Laplacian of a Gaussian (LoG) kernel with the perpendicular depth:

$$Blob = \nabla^2 Gauss(x, y, z) * depth_{\perp}[x, y] \tag{2}$$

Equation (2) can be made 2D by realizing that the z coordinate of the LoG kernel is equal to the perpendicular depth:

$$Blob = \nabla^2 Gauss(x, y, depth_{\perp}[x, y]) * depth_{\perp}[x, y] \tag{3}$$

Where

$$\nabla^2 Gauss(x, y, z) = \frac{\partial^2 Gauss}{\partial x^2} + \frac{\partial^2 Gauss}{\partial y^2} + \frac{\partial^2 Gauss}{\partial z^2} = \frac{1}{\sqrt{2\pi\sigma^2}^3}\left(\frac{d^2}{\sigma^4} - \frac{3}{\sigma^2}\right)e^{-d^2/2\sigma^2} \tag{4}$$

$$\sigma = radius/\sqrt{3} \tag{5}$$

$$d = \sqrt{(x - x_{center})^2 + (y - y_{center})^2 + (z - z_{center})^2} \tag{6}$$

The blob response can also be calculated as a difference of Gaussians (DoG) or determinant of the Hessian matrix, though the above method was found to run the quickest because it requires the least amount of computation per x, y coordinate. Figure 61 depicts the convolution kernel cut in half, the radius of the inner sphere is the radius specified. Red points have negative weight and green points have positive weight. All points shown have a weight greater than .05.

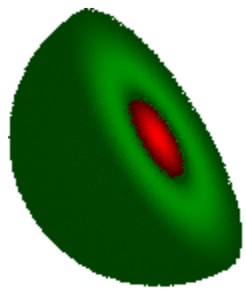
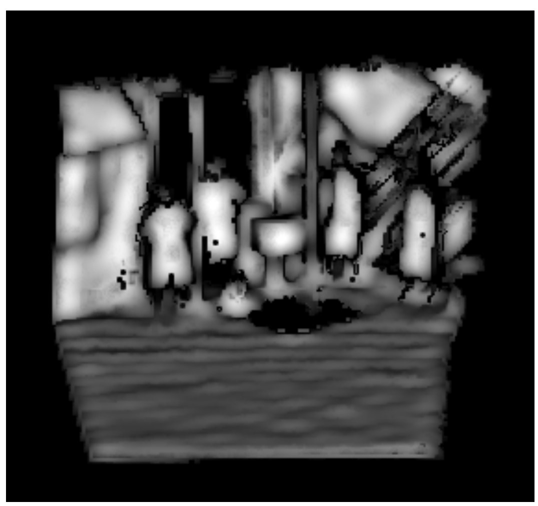


Figure 61: Cross section of the LoG kernel

Figure 62 shows the blob response when run with a radius of .4m and the resulting regions, created by grouping all of the pixels that are 4-connected. The final result narrows the image down to 25 regions of interest which is much more manageable than trying to classify every single point. The threshold value was chosen to be .6 because we've found that this works well.



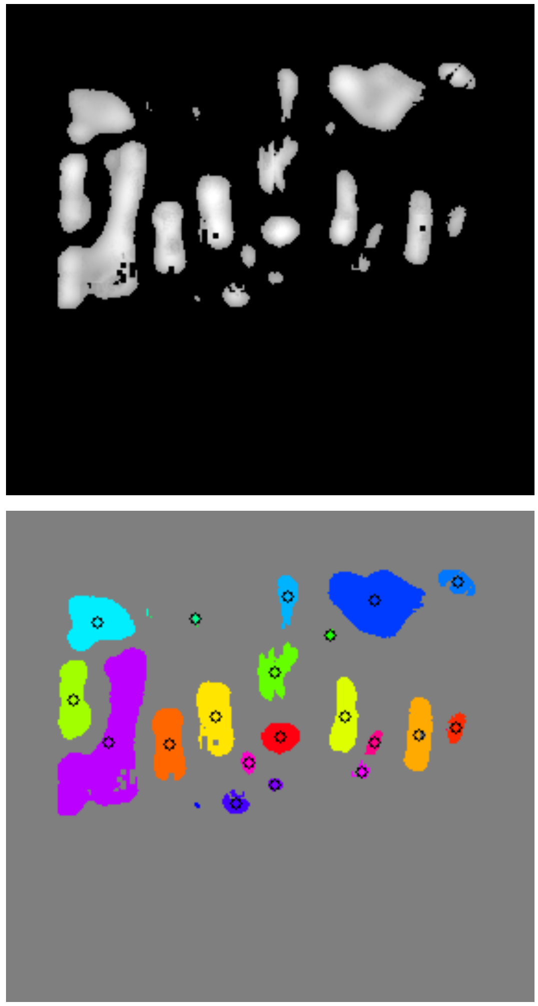


Figure 62: A- the normalized blob response with radius .4m B- the response with threshold of .6 applied C- the segmented regions with circles at the centroids of blobs with more than 10 points

Step 5: Generate structural tensors at each region of interest

The structural tensor is generated by finding the eigenvalues and eigenvectors of the covariance matrix of the gradients at each point convolved with some weight function, we simply use a Gaussian kernel.

$$S = \left(Weight * \begin{bmatrix} {g_x}^2 & g_x g_y & g_x g_z \\ g_y g_x & {g_y}^2 & g_y g_z \\ g_z g_x & g_z g_y & {g_z}^2 \end{bmatrix} \right) [x, y, depth_{\perp}(x, y)] \tag{7}$$

Where

$$Weight(x,y,z) = \begin{cases} 1, & (x,y) \in region\ of\ interest \\ Gauss(x,y,z), & otherwise \end{cases} \tag{8}$$

$$Gauss(x,y,z) = \frac{1}{\sqrt{2\pi * radius^2}^3} e^{-((x - x_{center})^2 + (y - y_{center})^2 + (z - z_{center})^2)/2radius^2} \tag{9}$$

$\langle g_x, g_y, g_z \rangle$ is found using equation 1 and the perpendicular depth:

$$g_x = tsdf(x - \Delta x, y, depth_\perp(x,y)) - tsdf(x + \Delta x, y, depth_\perp(x,y))$$

$$g_y = tsdf(x, y - \Delta y, depth_\perp(x,y)) - tsdf(x, y + \Delta y, depth_\perp(x,y))$$

$$g_z = tsdf(x, y, depth_\perp(x,y) - \Delta z) - tsdf(x, y, depth_\perp(x,y) + \Delta z) \tag{10}$$

$$S = U\Lambda U^T \tag{11}$$

Equation 11 is the eigensystem decomposition of S.

The eigenvectors of S define the semi-axes of the ellipse whereas the inverse of the corresponding eigenvalue determines the length. We need to take the inverse because the ellipse that bounds the region should be thinner where the gradient is large. The major axes are clamped to 6ft and 3ft to match the feature window while the length of the third axis is determined by its eigenvalue.

Currently we have implemented structural tensors in two ways: the user can hand select a region (Figure 63) or they can be generated using the regions of interest from step 4 (Figure 64). Selecting the regions by hand is useful for gathering training examples whereas the automatic generation will be used for classification. For now when we automatically create structural tensors we place them at the centroid of the region of interest and give each point within the region a weight of one, points that were not in the region of interest are weighted using a Gaussian kernel.

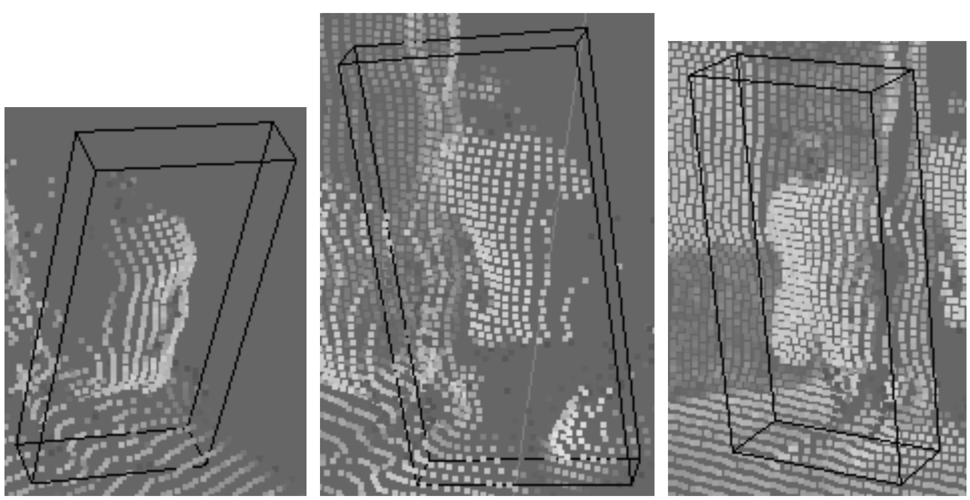


Figure 63: Some hand selected structural tensor regions

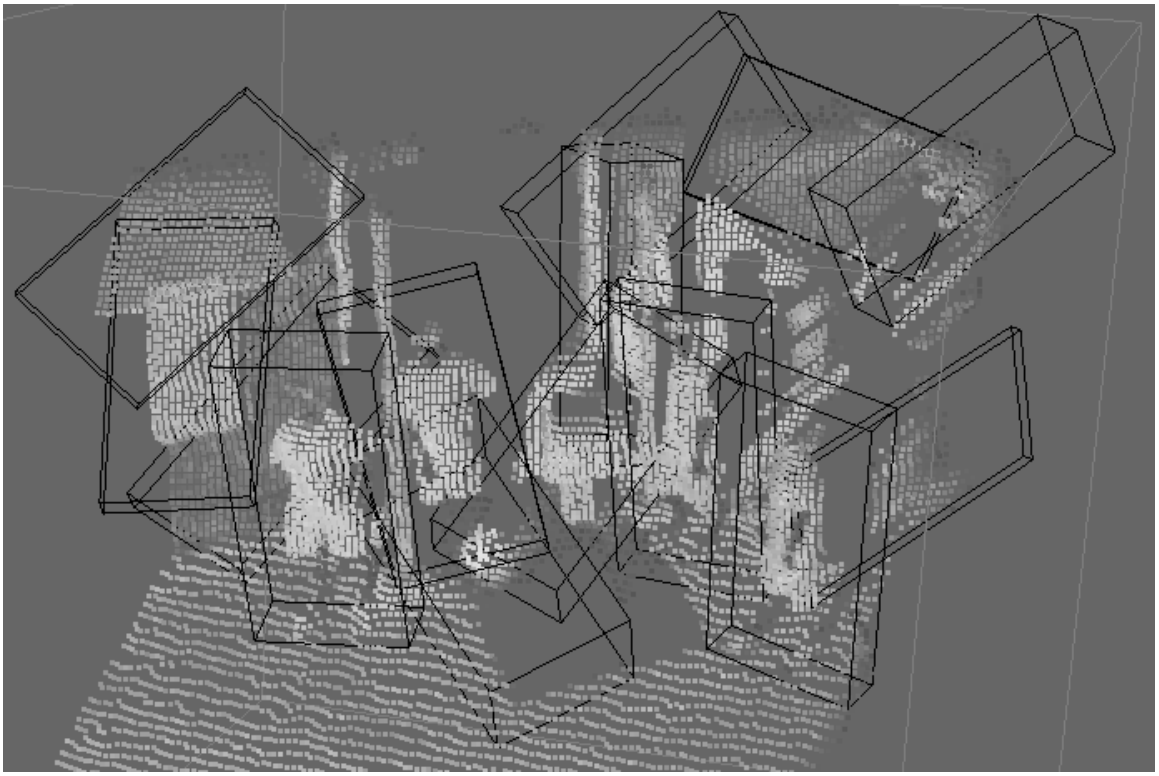


Figure 64: Automatically generated structural tensors using the regions from Figure 62 C

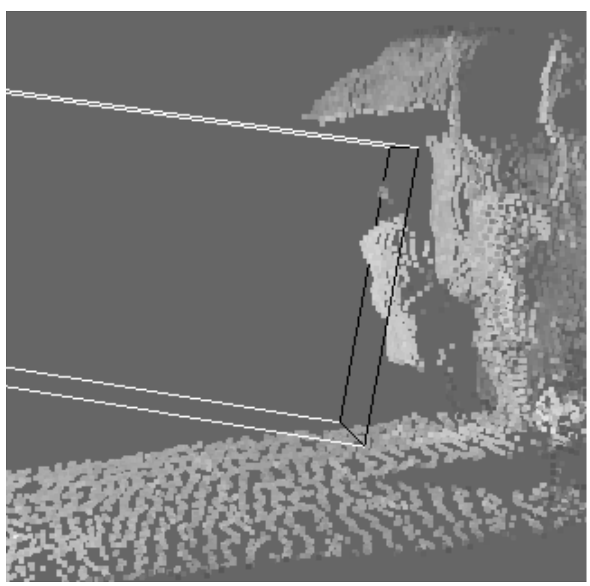


Figure 65: Poorly oriented window used in the sliding window approach

Step 6: Generate a feature image for each structural tensor

This step has not been implemented yet but the plan is to treat each structural tensor as its own TSDF volume and raycast the 20x40 depth image that we need for the training algorithm.

Step 7: Classify each image

Once we have a set of consistently oriented feature vectors they can be run through any training algorithm. While this step hasn't been done using the structural tensor images, we have trained and classified points in the depth by using an un-oriented sliding window approach. However, our scenes were shot from a balcony resulting in windows that poorly bounded the objects in the frame (compare Figure 63, Figure 64 and Figure 65). The sliding window was also only able to detect people that were standing straight up.

Step 8: Segment out the people from the original depth

At this point we have a set of structural tensors that have been classified as people. Each structural tensor can be thought of as a rectangular prism that bounds the person. Here we need to go back to using the original depth; for each valid depth we can calculate a 3D point and check to see if it lies within any of the structural tensors, if it does we assign a unique label based on which tensor it was contained in. After all the points have been checked we pass the segmented image to the tracker.

Implementation of 2-D and 3-D fusion algorithms

A particular focus during this period was fusion of image data across multiple modalities. Details of the algorithms are provided in [I.2.12], [I.2.30] and [I.2.26]. Image fusion across modalities is a challenging problem from accurate registration to meaningful representation of the fused information. A fused product must convey the important information from each modality in a way that can be naturally interpreted by a human observer. We are implementing a method of fusing 3D range information from a Flash LIDAR with a thermal MWIR image to convey the location of objects of interest within the focal plane and naturally within 3-space. The fusion method makes use of human visual perception of color and brightness to convey range and temperature, respectively.

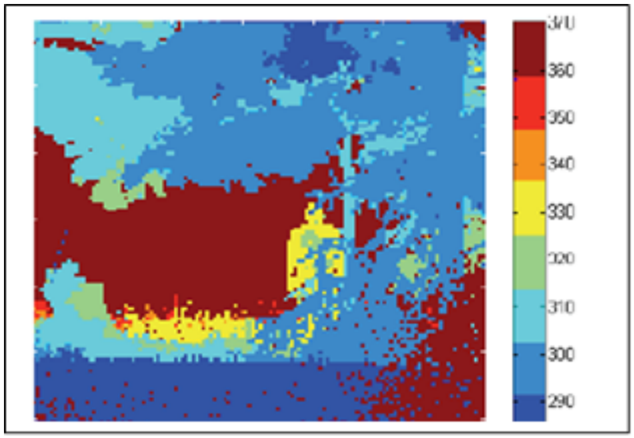


Figure 66: A flash LIDAR image mapped to 8 color bins within the ranges of 280 ft to 370 ft

Given a range image R directly mapped to ortho-rectified (x,y,z) points in 3-space, and a thermal image T, assumed to be registered pixel-by-pixel to the range image, a fused image may be constructed using data from each source. Given a color map, M_R with n bins, M_R: $R \rightarrow R_n$, divides the range image into n colors where $R_n \in \mathbb{R}^{px3}$ for p pixels (Figure 66). Each row of R_n provides a red, green, and blue value to define the pixel color. Given an intensity map M_I with k bins, $M_I \rightarrow I_k$ discretizes the intensity image into k bins, where $I_k \in \mathbb{R}^{px1}$ and each $I_k(i) \in [0,1]$, i=1,…,p. The fused image F is defined as $[I_k, I_k, I_k] \bullet M_R$, where $\bullet$ denotes the entry-wise product. This scales the intensity of the pixel colors by the intensity of the thermal image (Figure 67).

Figure 67: A MWIR thermal image mapped to 256 bins with values in [0 1]

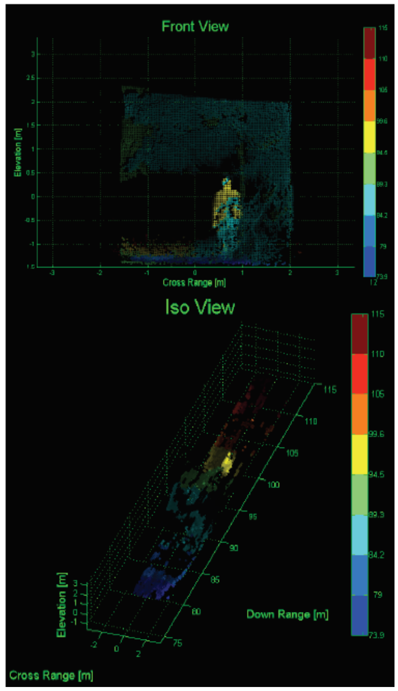
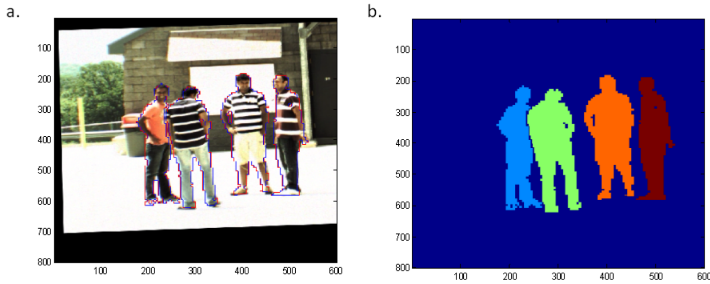


Figure 68: Two views of fused data. The human has a higher temperature than the background resulting in more vibrant colors that clearly indicate his range at 100m

Processing methods have also been developed for target classification applied to fusion of 3-D images and thermal images. The raw image data is converted to higher level information and a hierarchical object classification transforms the 3-D data into perceptual classes based on size shape and location. Subsequently the classification results are colored by point groupings, e.g., bushes, ground, pillars, roof, fascia and rails.

Range Image Tracking

A new tracking system has been developed that utilizes the Range Image or Depth Map from the Flash Lidar device. The steps of the algorithm are the following: i) detection and segmentation of regions, ii) evolution of the objects' contours (i.e., active contour, minimization of energy function, including motion filter), iii) linking the objects to tracks, and iv) adding new object information to a track database.



c.

ID
Sub ID
Area
Average Intensity
Occlusion
Active
MergeActive
Range to Target
Motion Statistics
State Vector x, y, z, x, y, z
Position, Velocity, Acceleration
Covariance Matrix
Kalman gain
Predicted Centroid (x,y,z)

d.

Frame	id	SubID	Centroid	Predicted Centroid	Area	MajorAxis Length	MinorAxis Length
1	1	1	411.8948 10.91759	411.8948 10.91759	5169	345.2527	31.95307
1	2	1	422.4857 210.1608	422.4857 210.1608	21895	417.9459	79.6264
1	3	1	392.2678 363.3222	392.2678 363.3222	19534	400.2308	73.39366
1	4	1	388.3873 418.4009	388.3873 418.4009	15701	413.3313	71.87224
1	5	1	386.1949 489.1579	386.1949 489.1579	18155	409.5271	62.02413
2	1	1	411.8948 10.91759	411.8948 10.91759	5169	345.2527	31.95307
2	2	1	422.4857 210.1608	422.4857 210.1608	21895	417.9459	79.6264
2	3	1	392.2678 363.3222	392.2678 363.3222	19534	400.2308	73.39366
2	4	1	388.3873 418.4009	388.3873 418.4009	15701	413.3313	71.87224
2	5	1	386.1949 489.1579	386.1949 489.1579	18155	409.5271	62.02413

Figure 69: Range Image Tracker Results: a. VNIR image with range segmented contours overlaid, b. range segmented image, c. tracker parameter output list, and d. tracker spreadsheet data base

The Interacting multiple Model (IMM) filter operates M Kalman filters in parallel, each of which is matched to a distinct motion model. It assumes that the transition between models is regulated by a finite-state Markov chain, with probability P_{ij} of switching from model I to model j in successive frame. However, rather than committing to any single model, it maintains a weighting among the models, which is determined as the probability of each model being correct given the current measurement. Hence, the optimal state estimate at any time instant is a mixture of Gaussian distributions. Each mixture component is the estimate from a Kalman filter, weighted by the posterior probability of the corresponding motion model.

This leads to a mixture with exponentially growing number of components in time because of branching of model switching hypothesis. To avoid the combinational explosion and make the computation tractable, the IMM filter approximates the mixture of Gaussian with a single Gaussian with equal mean and covariances. The probability of switch will later be used to give an indication of signaling events.

We are currently exploring using the motion filter along with the objects behavior to identify events. The motion models probabilities lend insight into the current motion that an object is undergoing. Monitoring the probabilities of the motion models combined with the trajectory can be used to signal potential events. For example, if a group of objects accelerate away from a common point (or accelerate towards a common point), it could be marked as an area of interest. Adding the processing of the body components in addition to the entire contour is another area we are considering. Not only would the segmentation and registration of the body components be needed, the relationship between the components will need to be modeled. In addition to providing a more accuracy determination of object pixels that are occluded, that capability would

provide both a macro level analysis of the scene and a micro level analysis of all the objects in the scene.

Particle Filter Color Tracker

Particle filters are usually used to estimate Bayesian models in which the latent variables are connected in a Markov chain similar to a hidden Markov model (HMM), but typically where the state space of the latent variables is continuous rather than discrete, and not sufficiently restricted to make exact inference tractable (as, for example, in a linear dynamical system, where the state space of the latent variables is restricted to Gaussian distributions and hence exact inference can be done efficiently using a Kalman filter). In the context of HMMs and related models, "filtering" refers to determining the distribution of a latent variable at a specific time, given all observations up to that time. Particle filters are so named because they allow for approximate "filtering" (in the sense just given) using a set of "particles" (differently-weighted samples of the distribution). Particle filters are the sequential ("on-line") analogue of Markov chain Monte Carlo (MCMC) batch methods and are often similar to importance sampling methods. Well-designed particle filters can often be much faster than MCMC. They are often an alternative to the Extended Kalman filter (EKF) or Unscented Kalman filter (UKF) with the advantage that, with sufficient samples, they approach the Bayesian optimal estimate. Thus, they can be made more accurate than either the EKF or UKF. However, when the simulated sample is not sufficiently large, they might suffer from sample impoverishment. The approaches can also be combined by using a version of the Kalman filter as a proposal distribution for the particle filter.

In general, tracking methods can be divided into two main classes specified as bottom-up or top-down approaches. In a bottom-up approach the image is segmented into objects which are then used for the tracking. For example blob detection can be used for the object extraction. In contrast, a top-down approach generates object hypotheses and tries to verify them using the image. Typically, model-based and template matching approaches belong to this class. The proposed color-based particle filter follows the top-down approaches, in the sense that the image content is only evaluated at the sample positions.

The proposed tracker employs the EMD distance to update the a priori distribution calculated by the particle filter. Each sample of the distribution represents an ellipse and is given as,

$$s = \left(x, y, \dot{x}, \dot{y}, H_x, H_y, \theta \right)$$

Where x, y, specify the location of the ellipse, $\dot{x}, \dot{y}$, the motion, H_x, H_y, the length of the half axes and θ the corresponding scale change and orientation. The tracker handles multiple hypotheses simultaneously. The sample set is propagated through the application of a dynamic model illustrated in Figure 70 and defined by

$$s_t = As_{t-1} + w_{t-1}$$

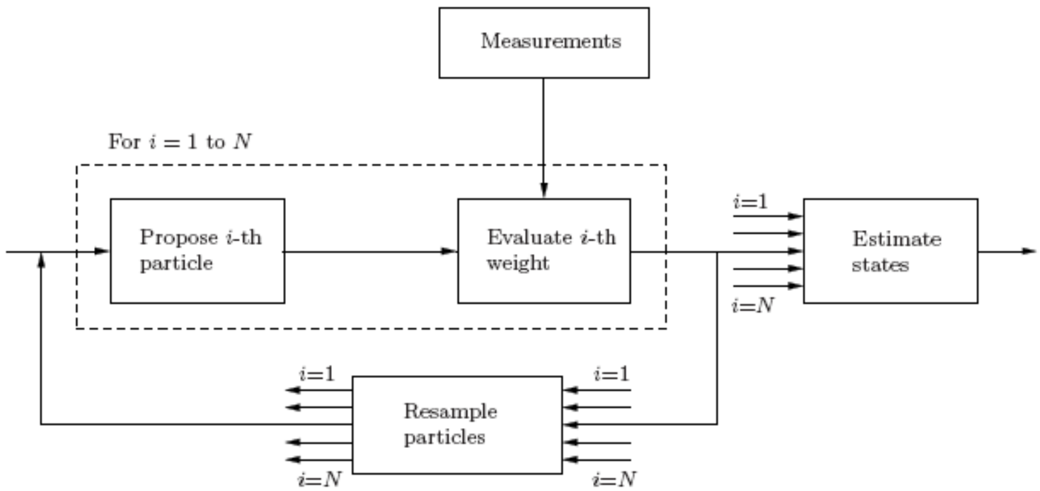


Figure 70: Particle Filter Block Diagram

The quantity A defines the deterministic component of the model and w_{t-1} is a multivariate Gaussian random variable. A first order model for A is used for describing a region moving with constant velocity $\dot{x}, \dot{y}$ and scale-change θ.

Figure 70 illustrates a typical tracker sequence.

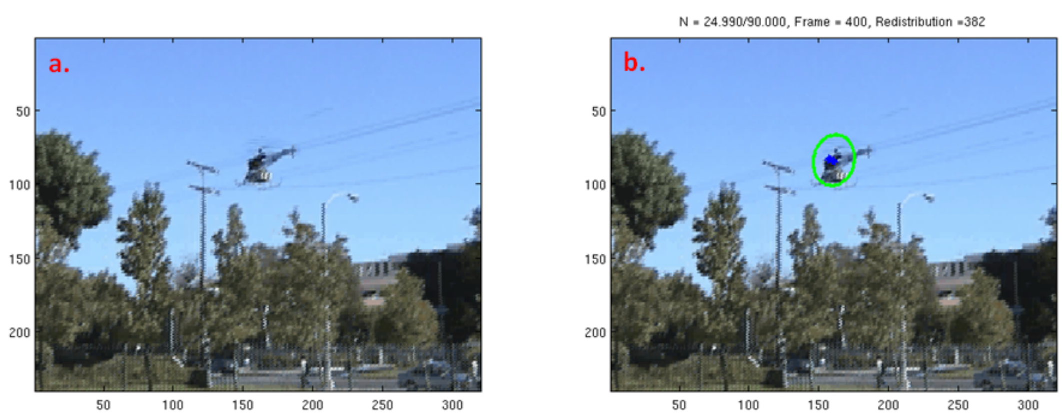


Figure 71: a. Original Image, b. Particle Filter Tracker Output

Figure 72 and Figure 73 show the particle filter tracker output "with" and "without" occlusion.

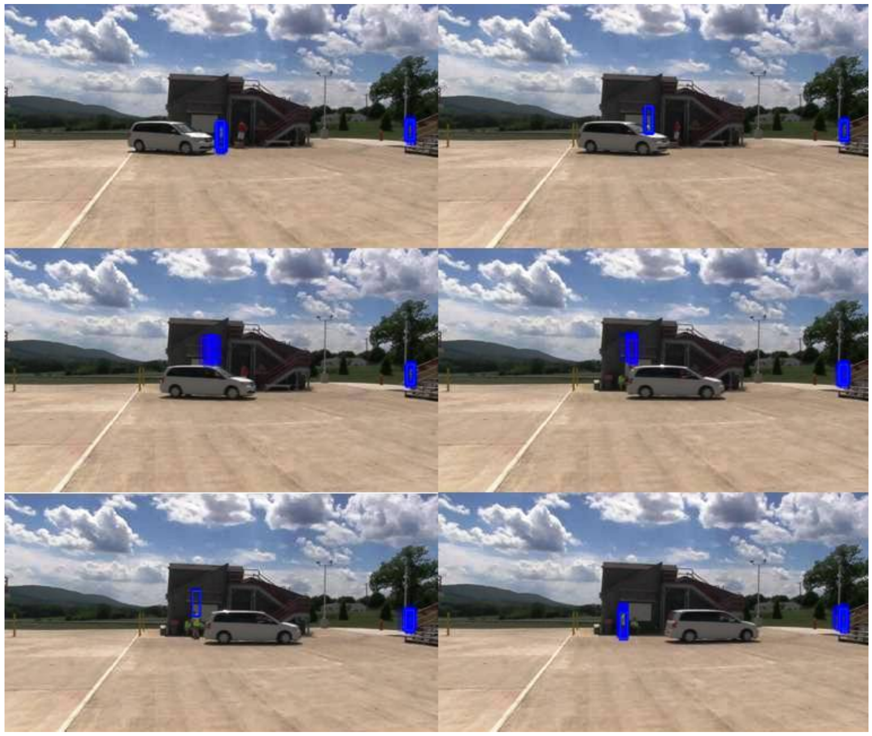


Figure 72: Particle Filter Tracker Output (Occlusion)

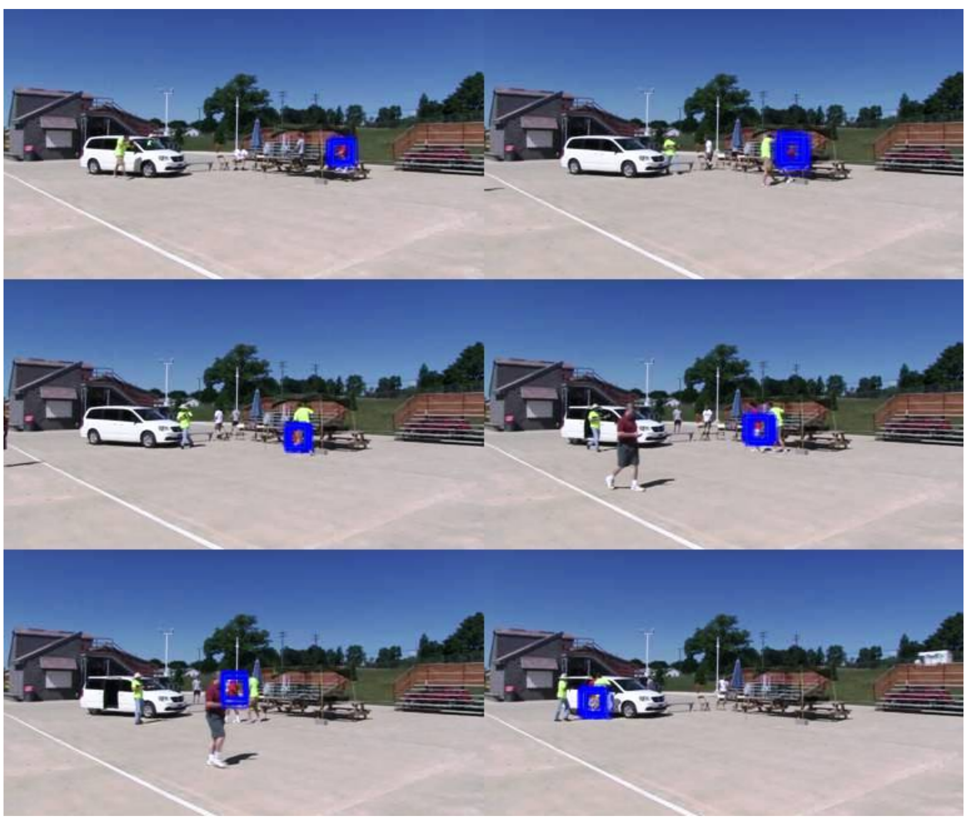


Figure 73: Particle Filter Output (No Occlusion)

Range - VINR Image Fusion

A particular focus during this period was fusion of image data across multiple modalities. Details of the algorithms are provided in [I.2.12], [I.2.30] and [I.2.26]. Image fusion across modalities is a challenging problem from accurate registration to meaningful representation of the fused information. A fused product must convey the important information from each modality in a way that can be naturally interpreted by a human observer. We are implementing a method of fusing 3D range information from a Flash LIDAR with a thermal MWIR image to convey the location of objects of interest within the focal plane and naturally within 3-space. The fusion method leverages concepts of human visual perception of color and brightness to convey range and temperature, respectively.

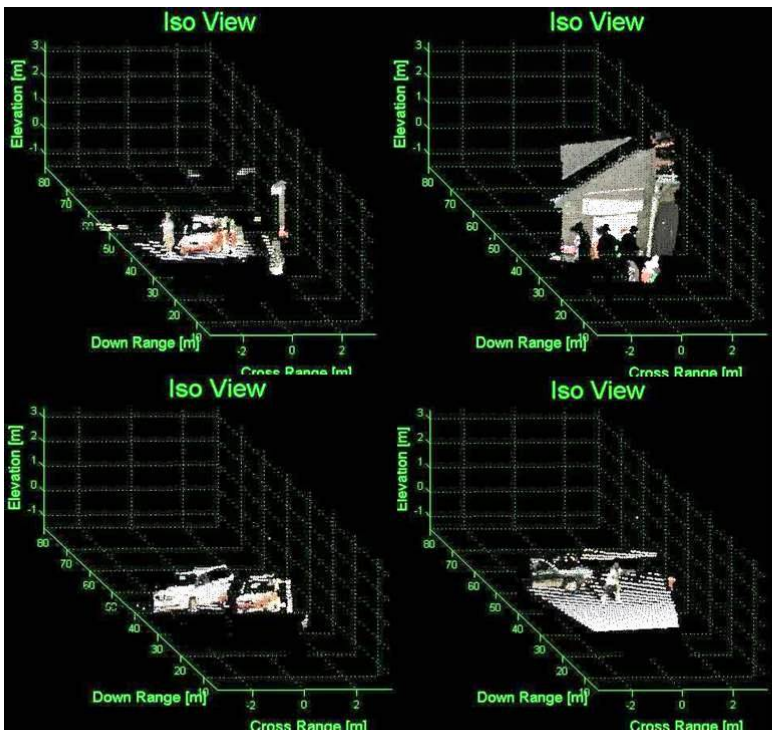


Figure 74: Range - VINR Image Fusion

Given a range image R directly mapped to ortho-rectified (x,y,z) points in 3-space, and a visible or thermal image T, assumed to be registered pixel-by-pixel to the range image, a fused image may be constructed using data from each source. Given a color map, M_R with n bins, M_R: $R \rightarrow R_n$, divides the range image into n colors where $R_n \in \mathbb{R}^{px3}$ for p pixels (Figure 66). Each row of R_n provides a red, green, and blue value to define the pixel color. Given an intensity map M_I with k bins, $M_I \rightarrow I_k$ discretizes the intensity image into k bins, where $I_k \in \mathbb{R}^{px1}$ and each I_k (i) € [0,1], i=1,…,p (Figure 67). The fused image F is defined as $[I_k, I_k, I_k] \bullet M_R$, where $\bullet$ denotes the entry-wise product. This scales the intensity of the pixel colors by the intensity of the thermal image. Processing methods have also been developed for target classification applied to fusion of 3-D images and thermal images. The raw image data is converted to higher level information and a hierarchical object classification transforms the 3-D data into perceptual classes based on

size shape and location. Subsequently the classification results are colored by point groupings, e.g., bushes, ground, pillars, roof, fascia and rails.

LWIR Tracker

This can serve as a baseline for the traditional approach of choosing Electro-Optic (EO) bands based on constraints imposed by the tracking scenario rather than opting for general hardware and trying to achieve specificity in software. It turns out that the tracker which was developed for use on EMD back-projected image sequences works naturally on LWIR imagery (The tracking is a much smaller computational load than the data preprocessing in the EMD tracker case). By using LWIR imagery one may skip heavy computational techniques entirely and use the thermal signature of the target in scalar form as input to the tracker. This thermal signature is achieved through device physics rather than computation. Below are some results, shown on the co-acquired color images in Figure 75.

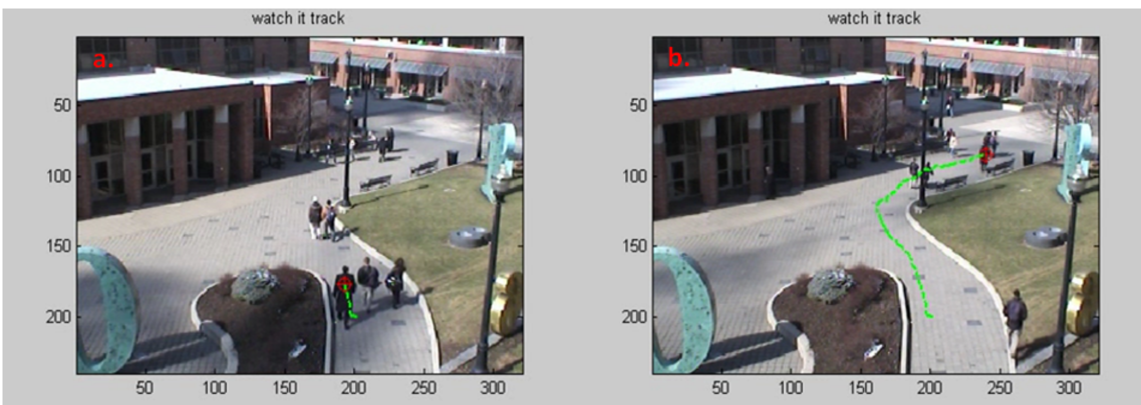


Figure 75: LWIR Tracker Local Hue Histogram Back-Projection Technique a. Initial Track, b. Final Track

The LWIR Tracker serves as a baseline depicting how the dismount tracking problem is typically solved. Rather than using multiple EO bands in some exotic way, one band is selected which suits the application. In the case of dismount tracking the thermal band is typically used. Sometimes a reduced subset of 2 or 3 bands can be used in a way which gives a strong target signature on skin, and then these bands can be fused into a single "skin" band which is directly analogous from the algorithm point of view to the thermal band as used in this example. By understanding how to constrain the system design given the details of a specific objective and selecting a single band or a fusion of a reduced subset of bands resulting in a single band we can use very straightforward and easy to implement processing techniques to accomplish robust tracking.

For implementation, MATLAB is a slow processing environment, particularly for data I/O and graphics. However as a testament to understanding and exploiting constraints posed by the end Vision objective, we can demonstrate greater than rea-time data I/O, processing, and visualization within the MATLAB environment. Since MATLAB is not data I/O and visualization optimized this is well below what is theoretically possible. The actual algorithm computation time within MATLAB functions at 4843 frames per second on average, roughly

161x real-time. This indicates that by implementing this end to end algorithm another platform which is data I/O and visualization real-time optimized, we can get extremely high frame rates on the order of several thousand frames per second with little effort, all through elegant design and professional understanding of EO/Vision system objectives.

Multi-Modal Detection of Man-Made Objects in Simulated Aerial Images

This approach uses a complex synthetic image generation model that produces simulated imagery in the visible through thermal infrared regions for multi-modal detection of man-made objects from aerial imagery. Detections are made in polarization imagery, hyper-spectral imagery, and LIDAR point clouds then fused into a single confidence map. The detections are based on reflective, spectral, and geometric features of man-made objects in airborne images. The polarization imagery detector uses the Stokes parameters and the degree of linear polarization to find highly polarizing objects. The hyper-spectral detector matches scene spectra to a library of man-made materials using a combination of the spectral gradient angle and the generalized likelihood ratio test. The LIDAR detector clusters 3D points into objects using principle component analysis and prunes the detections by size and shape. Once the three channels are mapped into detection images, the information can be fused without some of the problems of multi-modal fusion, such as edge reversal.

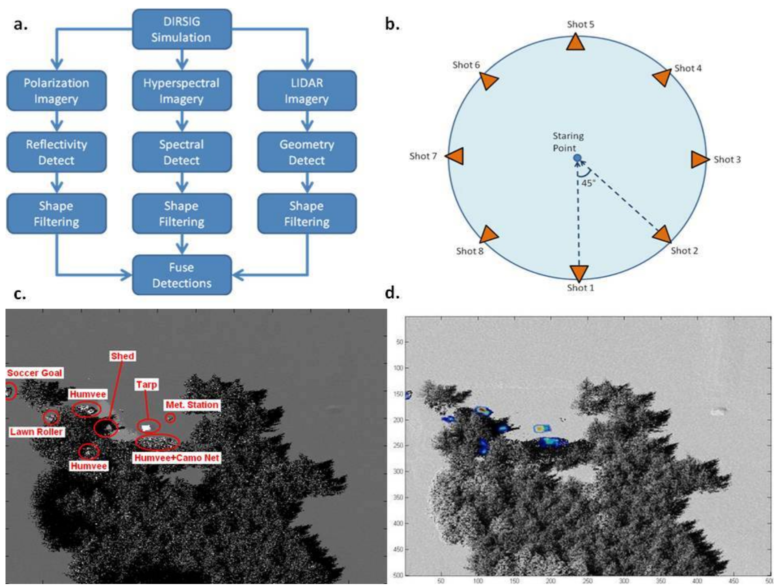


Figure 76: Multi-Sensor Fusion Process; a. Algorithm Flowchart, b. Simulated Flight Trajectory for Data Collect c. Ground Truth Object Map, d. Fused Sensor Detection heat Map

3.3.9 Penn State References

[I.1.1] D. Hall, J. Graham, M. McNeese, J. Rimland and R. Tutwiler, *Interim Progress report on the Army Research Office Multidisciplinary University Research Initiative (MURI) grant on Unified Research on Network-based Hard/Soft Information Fusion*, prepared for the U. S. Army Research Office, August, 1, 2011

[I.1.2] J. Graham, Scene *Setter for MURI Demonstration*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, July 30, 2012

[I.1.3] J. Graham et al. *Acronym List for SYNCOIN Data*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, December 2011

[I.1.4] J. Graham et al. *Identification and Location of Events/Activities associated with SYNCOIN Data*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, December 2011

[I.1.5] J. Graham et al. *Reference Maps for SYNCOIN Data*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, December 2011

[I.1.6] J. Rimland, *Database Schema for SYNCOIN Data*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, December 2011

[I.1.7] J. Graham et al, *Analyst Notebook Charts*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802, May 1, 2012

[I.1.8] Geosuite website; http://www.gdc4s.com/, downloaded on August 5, 2012

[I.1.9] Centre County Public Safety Training facility; https://sss.cpi.edu/est, downloaded on August 5, 2012

[I.1.10] Hall, D., McNeese. M., Llinas, J., and Mullen, T., "A framework for hard/soft fusion," *Proceedings of the 11th International Conference on Information Fusion,* Cologne, Germany, June 30 – July 03, 2008

[I.1.11] Buchmann, A and B Koldehofe. 2009. Complex event processing. *It-Information Technology* 51 (5): 241-242.

[I.1.12] Taylor, K and L Leidinger. 2011. Ontology-driven complex event processing in heterogeneous sensor networks. *The Semanic Web: Research and Applications,* 285-299

[I.1.13] Wagner, A, D Anicic, R Stühmer, N Stojanovic, A Harth, and R Studer. Linked data and complex event processing for the smart energy grid. In *Proc. 2010 of CEUR Workshop, ISSN*

[I.1.14] Hu, W, W Ye, Y Huang, and S Zhang. Complex Event Processing in RFID Middleware: A Three Layer Perspective. In Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on. 2008

[I.1.15] Museux, N, J Mattioli, C Laudy, and H Soubaras. "Complex Event Processing Approach for Strategic Intelligence." In Information Fusion, 2006 9th International Conference on. 2006

[I.1.16] Wooldridge, M. 2002. *An Introduction to Multiagent Systems.* Wiley

[I.1.17] Castanedo, F, J García, M A Patricio, and J M Molina. 2008. Analysis of distributed fusion alternatives in coordinated vision agents. In *Information Fusion, 2008 11th International Conference on*

[I.1.18] Fan, X, M McNeese, B Sun, T Hanratty, L Allender, and J Yen. 2010. Human--agent collaboration for time-stressed multicontext decision making. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 40 (2): 306-320

[I.1.19] Fan, X, S Sun, M McNeese, and J Yen. 2005. Extending the recognition-primed decision model to support human-agent collaboration. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems.*

[I.1.20] Keisler, R J. 2008. Towards an Agent-Based, Autonomous Tactical System for C4ISR Operations SPACE AND NAVAL WARFARE SYSTEMS CENTER NORTH CHARLESTON SC

[I.1.21] Logenthiran, T, D Srinivasan, and A M Khambadkone. 2011. Multi-agent system for energy resource scheduling of integrated microgrids in a distributed system. *Electric Power Systems Research* 81 (1): 138-148

[I.1.22] Olfati-Saber, R. 2006. Flocking for multi-agent dynamic systems: Algorithms and theory. *Automatic Control, IEEE Transactions on* 51 (3): 401

[I.1.23] Schoenharl, T, G Madey, G Szabó, and A L Barabási. 2006. WIPER: A multi-agent system for emergency response. In *Proceedings of the 3rd International ISCRAM Conference*

[I.1.24] D. J. Cook and S. K. Das, "How smart are our environments? An updated look at the state of the art," Pervasive Mob. Comput., vol. 3, no. 2, pp. 53–73, Mar. 2007.

3.3.10 Publications

[I.2.1] J. Rimland and M. Ballora, (2014), "Using Complex Event Processing (CEP) and vocal synthesis techniques to improve comprehension of sonified human-centric data," *Proceedings of SPIE 2014*, Baltimore, MD, May 6, 2014.

[I.2.2] J. Rimland and M. Ballora, (2014), "Using vocal-based sounds to represent sentiment in complex event processing", *Proceedings of the International Conference on Auditory Display (ICAD),* 2014

[I.2.3] J. Rimland, S. Shaffer and D. L. Hall, (2014) "A hitchhiker's guide to distributed hard and soft information fusion infrastructure development", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain.

[I.2.4] S. Shaffer, (2014), "Automatic theory generation from analyst text files using coherence networks", *Proceedings of the SPIE 2014 Conference, Baltimore*, MD, May 6, 2014.

[I.2.5] G. Cai and J. Graham (2014), "Semantic data fusion through visually-enabled analytical reasoning", *Proceedings of International Society of Information Fusion FUSION 2014*, July, 2014, Salamanca, Spain

[I.2.6] G. Cai, G. Gross, J. Llinas and D. Hall (2014), "A visual analytic framework for data fusion in investigative intelligence", *Proceedings of the SPIE volume 9122, Next-Generation Analyst II*. Baltimore, Maryland, USA\

[I.2.7] John P. Morgan and Richard L. Tutwiler, " Real-Time reconstruction of depth sequences using signed distance functions ", *Proceedings of the SPIE 2014 Conference*, Baltimore, MD, May 6, 2014

[I.2.8] Sherry, R., J. Gabor and D. Hall (2009), "Information fusion to support real time accident diagnosis and accident management", *Proceedings of the OECD/NEA Workshop on Implementation of Severe Accident Management (SAM) Measures*, October 26-28, 2009, Schloss Bottstein, Switzerland

[I.2.9] Hall, D., Hellar, B., and McNeese, M. D., (2009), "The Extreme Events Laboratory: A cyber infrastructure for performing experiments to quantify the effectiveness of human-centered information fusion," *Proceedings of the 2009 International Conference on Information Fusion (Fusion 2009),* Seattle, Washington, July, 2009

[I.2.10] D. Hall, J. Graham, L. More and J. Rimland (2010), "Test and evaluation of soft/hard information fusion systems: an experimental environment, methodology and initial data sets", *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, UK, July, 2010

[I.2.11] N. Giacobe, (2010), "Mining social media in extreme events: lessons learned from the DARPA network challenge", *Proceedings of the IEEE Conference on Homeland Securities Technologies (IEEE HST 2010),* Waltham, MA, November 2010

[I.2.12] J. Llinas, R. Nagi, D. Hall and J. Lavery (2010), "A multidisciplinary university research initiative in hard and soft information fusion: overview, research strategies and initial results," in *Proceedings of the 13th International Conference on Information Fusion*, Edinburgh, UK, July, 2010

[I.2.13] R. Tutwiler, D. J. Natale, M. S. Baran, R. L. Tutwiler, (2010), "Live motion 3D data processing", *Proceedings of the IDGA 9th Image Fusion Summit,* November 15 - 17, 2010, Sheraton Premiere at Tysons Corner, Vienna, VA

[I.2.14] J. Graham, J. Rimland, D. Hall, (2011), "A COIN-inspired synthetic data set for quantitative evaluation of hard and soft fusion systems", *Proceedings of Fusion 2011: the International Conference on Information Fusion,* Chicago, IL, July, 2011

[I.2.15] R. Tutwiler, M. Baran, D. Natale, C. Griffin, J. Daughtry, M. McQuillan, J. Rimland, and D. Hall, (2011), "Hard sensor fusion for COIN inspired situation awareness",

Proceedings of Fusion 2011: the International Conference on Information Fusion, Chicago, IL, July, 2011

[I.2.16] J. Rimland, (2011), "A multi-agent infrastructure for hard and soft information fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011

[I.2.17] J. Rimland, (2011), "JDL level 0 and 1 algorithms for processing and fusion of hard sensor data", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011

[I.2.18] J. Graham (2011), "A new synthetic dataset for evaluating soft and hard fusion algorithms", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011

[I.2.19] D. J. Natale, M. S. Baran, R. Tutwiler and D. L. Hall, (2011), "3DSF: three dimensional spatio-temporal fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25-29 April, 2011

[I.2.20] D. Hall, (2011) invited panel discussion, "Real world issues and challenges in hard and soft data fusion", *Proceedings of the SPIE Defense, Security, and Sensing Symposium: Defense Transformation and Net-Centric Systems 2011*, Orlando, FL, 25 April, 2011

[I.2.21] J. Graham, J. Rimland and D. Hall (2011), "A COIN-inspired synthetic data set for qualitative evaluation of hard and soft fusion systems", *Proceedings of the 14th International Conference on Information Fusion*, Chicago, IL, July, 2011

[I.2.22] D. Hall, G. Iyer, M. Ballora, R. Cole, H. Kruesi and H. Greene, (2011), "Use of auditory displays in anomaly detection, *Proceedings of the National Symposium on Sensor and Data Fusion*, Oct. 24-28, 2011

[I.2.23] D. Hall, G. Iyer, M. Ballora, R. Cole, H. Kruesi and H. Greene, (2011), "Use of auditory displays in anomaly detection, *Proceedings of the National Symposium on Sensor and Data Fusion*, Oct. 24-28, 2011

[I.2.24] J. Rimland, D. Hall and J. Graham, (2012), "Human cognitive and perceptual factors on JDL level-4 hard/soft fusion", *Proceedings of the SPIE Conference on Multi-sensor, Multisource Information Fusion: Architectures, Algorithms*, and Applications 2012, Baltimore, MD, April 23-27, 2012

[I.2.25] D. Sudit, S. Kumara and D. Hall, (2012), "Complex mathematical model for soft processes in information fusion," *Proceedings of the ISERC 2012 Conference*, Orlando, FL, April, 2012

[I.2.26] J. Graham and D. Hall, (2012), "The use of Analytic Decision Game (ADG) methods for test and evaluation of hard and soft data fusion systems and education of a new

generation of data fusion analysts," *Proceedings of the National Symposium on Sensor Data Fusion (NSSDF),* Washington, DC, 22-16 October, 2012

[I.2.27] D. Kretz, B. Simpson and J. Graham, (2012), "A Game-Based Experimental Protocol for Identifying and Overcoming Judgment Biases in Forensic Decision Analysis", *IEEE International Conference on Technologies for Homeland Security*, Waltham, MA, 13-15 November, 2012

[I.2.28] Matthew S. Baran; Richard L. Tutwiler; David L. Hall; Donald J. Natale "Surface reconstruction for 3D remote sensing", *Proceedings of SPIE 2012*, Baltimore, MD

[I.2.29] J. Rimland, D. Coughlin, D. Hall and J. Graham, (2012),"Advances in data representation for hard/soft information fusion", *Proceedings of SPIE 2012*, Baltimore, MD

[I.2.30] J. Rimland and J. Llinas, (2012), "Network and infrastructure considerations for hard and soft information fusion processes,", *Proceedings of the International Society of Information Fusion*, FUSION 2012, July, 2012, Singapore

[I.2.31] J. Rimland, M. Ballora and D. Hall, (2013), "Hard and soft information fusion in sonification for assistive mobile device technology", *Proceedings of the International Conference on Auditory Display* (ICAD_2013), July 6 – 10, 2013, Lodz University of Technology, Poland

[I.2.32] M.S. Baran, R.L. Tutwiler, D. J. Natale, M .S. Bassett, M. P. Haner, (2013), "Multi-Modal Detection of Man-Made Objects in Simulated Aerial Imagery", *Proc. SPIE*. 8743*, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultra-spectral Imagery XIX* 87430P (May 18, 2013)

[I.2.33] N. S. Butler, R.L. Tutwiler, (2013), "Invariant unsupervised segmentation of dismounts in depth images", *Proc. SPIE*. 8745, *Signal Processing, Sensor Fusion, and Target Recognition* XXII 87451B (May 23, 2013)

[I.2.34] J. Rimland and M. Ballora ,(2013), "Beyond visualization of Big Data: a multi-stage data exploration approach using visualization, sonification, and storification", *Proceedings of SPIE 2013*.

[I.2.35] J. Rimland, M. McNeese and D. Hall, (2013), "Conserving Analyst Attention Units: Use of Multi-agent Software and CEP Methods to Assist Information Analysis", *Proceedings of SPIE DSS Conference on Next-Generation Analyst, vol. 8758, April, 2013, Baltimore, Md.*

3.3.11 Books and Book Chapters

[I.3.1] D. Hall, J. Llinas, C. Chong, K. C. Chang, editors, *Distributed Data Fusion for Network-Centric Operations,* CRC Press, August, 2012

[I.3.2] D. L. Hall, "Perspectives on Distributed Data Fusion", chapter 1 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press*,* August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang

[I.3.3] J. Rimland, "Service-Oriented Architecture for Human-Centric Information Fusion," chapter 13 in *Distributed Data Fusion for Network-Centric Operations,* CRC Press*,* August, 2012, edited by D. Hall, J. Llinas, C. Chong and K. C. Chang

[I.3.4] D. Hall, "The Emergence of Human-Centric Information Fusion," chapter 27 in *Distributed Sensor Networks*, 2nd edition, 2012, edited by S. Iyengar and R. Brooks

[I.3.5] Hall, D. and Jordan, J. (2010). *Human Centered Information Fusion*. Artech House, Inc.

[I.3.6] D. Hall and S. Aungst (2010), The use of soft sensors and I-Space for improved combat ID, chapter 10 in *Human Factors in Combat Identification*, ed. by D. Andrews, R. Herz and M. Wolf, Ashgate, pp 161-170

[I.3.7] D. Hall, (2012), "Understanding the new users: collaborative decision-making paradigms, communities of interest, and complex adaptive systems", chapter 3 in D. Hall, J. Llinas, C. Chong, K. C. Chang, editors, *Distributed Data Fusion for Network-Centric Operations,* CRC Press, 2012

[I.3.8] D. L. Hall, C. M. Hall, S. A. H. McMullen and M. McMullen (2011), "Improving uncertainty assessment and situational awareness using hard and soft information fusion", in *Risk Management in Decision Making: Intelligent Methodologies and Applications,* edited by J. Lu, Springer, 2011

[I.3.9] D. Hall, (2011), "The Emergence of Human-Centric Information Fusion," chapter 18 in *Distributed Sensor Networks*, 2nd edition, 2011

[I.3.10] D. Hall, (2012), "The Emergence of Human-Centric Information Fusion," chapter 27 in Distributed Sensor Networks, 2nd edition, 2012, edited by S. Iyengar and R. Brooks

3.3.12 Technical Reports

[I.4.1] Saab and F. Fonseca, (2009), *Participatory Sensing: A Review of the Literature and State of the Art Practices,* Technical Report for the Penn State University Center for Network Centric Cognition and Information Fusion (NC2IF), November 11, 2009 (78 pages)

[I.4.2] D. L. Hall and M. McNeese (2010), *First year interim report for the Multidisciplinary University Research Initiative (MURI) on Unified Research on Network-based Hard/Soft Information Fusion*, prepared for the U. S. Army Research Office, August, 23, 2010

[I.4.3] R. L. Tutwiler, *MURI Hard Sensor Fusion Performance Characterization*, Technical report, May, 2011

[I.4.4] J. Rimland, (2011) "Factors determining success in participatory sensing campaigns.", Internal Report for the NC^2IF Research Center, January, 2011

[I.4.5] J. Rimland, (2011), "Cognitive factors in data fusion and visualization", Internal Report for the NC^2IF Research Center, March, 2011

[I.4.6] J. Rimland, (2011), "The role of perceptual factors in human-in-the-loop HCI", Internal Report for the NC^2IF Research Center, May, 2011

[I.4.7] D. Hall, J. Graham, M. McNeese, J. Rimland and R. Tutwiler (2011), Second Year Interim Progress Report: *Army Research Office Multidisciplinary University Research Initiative (MURI) grant on Unified Research on Network-based Hard/Soft information Fusion*, August 23, 2011

[I.4.8] D. L. Hall, J. Graham, M. McNeese, J. Rimland and R. Tutwiler, (2012) , *Third Year Interim Progress Report: Army Research Office Multidisciplinary University Research Initiative (MURI) grant on Unified Research on Network-based Hard/Soft information Fusion*, August 23, 2012 (28 pages)

[I.4.9] R. L. Tutwiler, *MURI Hard Sensor Fusion Performance Characterization*, Technical report, May, 2011

[I.4.10] J. Graham, *SYNCOIN Data Set*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC^2IF) Research Center, IST Building, University Park, PA 16802, revised, December, 2011

[I.4.11] N. Giacobe, *SYNCOIN Word Clouds*, Technical report prepared for the Penn State Network Centric Cognition and Information Fusion (NC2IF) Research Center, IST Building, University Park, PA 16802 May 1, 2012

[I.4.12] D. L. Hall, J. Graham, M. McNeese, J. Rimland, R. Tutwiler and G. Cai, (2013), *Unified Research on Network-based Hard/Soft Information Fusion*, Interim progress report for the Army Research Office Multidisciplinary University Research Initiative, July, 2013, (42 pages)

[I.4.13] J. Graham et al, (2014), *Analyst Workbench Instructional Guide,* Technical report for the NC^2IF Research Center, August, 2014 (30 pages)

3.3.13 Theses and Dissertations

[I.5.1] S. R. Nimmala (2014), *Architectural considerations for context aware applications in mobile cloud computing environment*, M.S. thesis in Computer Science and Engineering, The Pennsylvania State University, August 2014

[I.5.2] K. Misra, (2010) A cyber infrastructure for hard and soft data fusion, M. S. thesis in Electrical Engineering, The Pennsylvania State University, University Park, PA

[I.5.3] Xu M.S. (2010) Unsupervised flow-level clustering in network anomaly detection, M. S. thesis in Electrical Engineering, The Pennsylvania State University, University Park, PA

[I.5.4] Rachana Reddy Agumamidi, (2011) ,M. S. thesis, The Pennsylvania State University, Electrical Engineering, “Hard Sensor Processing for Data Fusion”, May, 2011

[I.5.5] Ganesh Iyer, (2011), M.S. thesis, The Pennsylvania State University, Electrical Engineering, “Approaches to hard and soft sensors’ data fusion”, June, 2011

[I.5.6] A. Godbole, (2013) Improving utilization of mobile device technology for distributed emergency teams, M.S. theses in Computer Science and Engineering, The Pennsylvania State University, June 2013

[I.5.7] J. C. Rimland (2013), Hybrid human-computing distributed sense-making: Extending the SOA paradigm for dynamic adjudication and optimization of human and computer roles”, Ph.D. dissertation in Information Sciences and Technology, The Pennsylvania State University, August, 2013

3.4 Tennessee State University

3.4.1 Tennessee State Abstract

This report presents a summary of the research activities performed by the Tennessee State University (TSU) during the period of 2009-2014 on the US Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) project entitled “Unified Research on Network-based Hard/Soft Information Fusion”. This project is a three-University project led by University at Buffalo (UB), and participated by the Pennsylvania State University (PSU) and Tennessee State University (TSU). TSU research responsibilities on this project included the followings:

(1) Development of suitable taxonomy and ontology for recognition of human-vehicle interactions (HVI), human-human interactions (HHI), and human-object interactions (HOI).
(2) Development of robust architectural framework with appropriate supportive computational models and techniques for multi-modality hard sensor fusion.
(3) Conduct human-in-the-loop experiments for characterization and discovery of suspicious social networks and group activities based on the capabilities of newly developed architectural framework for multi-modality sensor fusion.
(4) Develop a method for attribute-based characterization, and semantic annotation of sensors observed social networks and group activities;
(5) Test and validate the efficiency and effectiveness of newly developed multi-modality sensor fusion techniques and algorithms.
(6) Support the integration of hard/soft sensor fusion effort of University of Buffalo in collaboration with the Penn State University.
(7) Transition of hard sensor processing techniques to ARL.

This report is organized in two parts: (1) the project related statistics, and (2) project research accomplishments summary for the period of this project.

3.4.2 Tennessee State Project Related Statistics

List of papers submitted or published:

j) Papers published in peer-reviewed journals - N/A

k) Papers published in non-peer-reviewed journals or in conference proceedings:

1. Alkilani, A., and Shirkhodaie, A., "Acoustic Events Semantic Detection, Classification, and Annotation for Persistent Surveillance Applications", SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.
2. Elangovan, V., and Shirkhodaie, A., "Knowledge Discovery in Group Activities Through Sequential Observation Analysis", SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.
3. Habibi, M. S., and Shirkhodaie, A., "Multi-attributed Tagged Big Data Exploitation for Hidden Concepts Discovery", SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.
4. Shirkhodaie, A., Elangovan, V., Habibi, M. S., and Alkilani, A., "A Decision Support System for Fusion of Soft and Hard-sensor Information Based on Latent Semantic Analysis Technique", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
5. Elangovan, V., and Shirkhodaie, A., "A Robust Technique for Group Activities Recognition Based on Fusion of Extracted Features in Video Streams", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
6. Elangovan, V., Bashir, A., and Shirkhodaie, A., "A Multi-attribute Based Methodology for Vehicle Detection & Identification", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
7. Elangovan, V., Alkilani, A., and Shirkhodaie, A., "A Multi-Modality Attributes Unmasking Scheme for Group Activity Characterization and Data Fusion", IEEE Intelligence and Security Informatics (ISI), Seattle, WA, June 2013.
8. Alkilani, A., and Shirkhodaie, A., "Acoustic Recognition of Human-Object Interactions in Persistent Surveillance Systems", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
9. Habibi, M. S., and Shirkhodaie, A., "Mining Patterns in Persistent Surveillance Systems with Smart Query and Visual Analytics", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
10. Vinayak Elangovan, Amir Shirkhodaie, "Team Activity Analysis and Recognition Based on Kinect Depth Map and Optical Imagery Techniques*", SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-30, April 23, 2012, Baltimore, MD.*
11. Vinayak Elangovan, Amir Shirkhodaie, "Recognition of Human Activity Characteristics Based on State Transitions Modeling Technique*", SPIE 2012 Defense, Security, and*

Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-43, April 23, 2012, Baltimore, MD.*

12. Amjad Alkilani, and Amir Shirkhodaie, "A Survey on Acoustic Signature Recognition and Classification Techniques for Persistent Surveillance Systems*", SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-28, April 23, 2012, Baltimore, MD.*
13. *Mohammad Habibi, and Amir Shirkhodaie, "A Survey of Visual Analytics for Knowledge Discovery and Content Analysis," SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-27, April 23, 2012, Baltimore, MD.*
14. Vinayak Elangovan, Amir Shirkhodaie, "Human Activity Discovery & Recognition based on State Transitions Modeling in Persistent Surveillance Systems," SPIE Electronic Imaging Science and Technology Conference, Multimedia Content Access: Algorithms and Systems VI, 22 - 26 January 2012, San Francisco, CA.
15. Vinayak Elangovan, Amir Shirkhodaie, "Adaptive Characterization, Tracking, and Semantic Labeling of Human-Vehicle Interactions via Multi-modality Data Fusion Model," SPIE Electronic Imaging Science and Technology Conference, Multimedia Content Access: Algorithms and Systems VI, 22 - 26 January 2012, San Francisco, CA.
16. Amir Shirkhodaie, "Perceptual Semantic Labeling of Human-Vehicle Interactions (HVI)," 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2012, Baltimore, MD.
17. Amir Shirkhodaie, "Semantic Labeling of Human-Vehicle Interactions Via Acoustic Events Characterization and Inference," 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2012, Baltimore, MD.
18. Vinayak Elangovan, Haroun Rababaah, Amir Shirkhodaie, "Acoustic Semantic Labeling and Fusion of Human-Vehicle Interactions," SPIE Defense, Security, and Sensing Conference, Signal Processing, Sensor Fusion and Target Recognition XX, April 25, 2011, Orlando, FL.
19. Vinayak Elangovan, Amir Shirkhodaie, "Context-based semantic labeling of human-vehicle interactions in persistent surveillance systems," SPIE Defense, Security, and Sensing Conference, Visual Information Processing XX, April 26, 2011, Orlando, FL.
20. Vinayak Elangovan, Amir Shirkhodaie, "A survey of imagery techniques for semantic labeling of human-vehicle interactions in persistent surveillance systems," SPIE Defense, Security, and Sensing Conference, Signal Processing, Sensor Fusion, and Target Recognition XX, April 25, 2011, Orlando, FL.
21. Vinayak, E., Shirkhodaie, A., and Rababaah, A., "Context-Based Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems, 13th International Conference on Information Fusion, 26-29 July 2010 EICC Edinburgh, UK.

22. Rababaah, H., Shirkhodaie, A., “Semantic Labeling of Non-Stationary Vehicular Acoustic Events in Persistent Surveillance Systems,” 13th International Conference on Information Fusion, 26-29 July 2010 EICC Edinburgh, UK.
23. Rababaah, H., Shirkhodaie, A., "Twitter web-service for soft agent reporting in persistent surveillance systems" in Cyber Security, Situation Management, and Impact Assessment II; and Visual Analytics for Homeland Defense and Security II, edited by John F. Buford, Gabriel Jakobson, John Erickson, Proceedings of SPIE Vol. 7709, 2010.
24. Shirkhodaie, A., Rababaah, H., "Multi-layered context impact modulation for enhanced focus of attention of situational awareness in persistent surveillance systems" in Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010, edited by Jerome J. Braun, Proceedings of SPIE Vol. 7710, 2010.

Presentations

a) Non-peer reviewed journals – N/A

2009 Presentations:

1. ARO-MURI Kick-Off Meeting, University of Buffalo, Buffalo, NY, August 25, 2009.

2010 Presentations:

1. Rababaah, H., Shirkhodaie, A., "Twitter Web-Service for Soft Agent Reporting in Persistent Surveillance Systems" in Cyber Security, Situation Management, and edited by John F. Buford, Gabriel Jakobson, John Erickson, Proceedings of SPIE Vol. 7709, 2010.
2. Shirkhodaie, A., Rababaah, H., "Multi-layered Context Impact Modulation for Enhanced Focus of Attention of Situational Awareness in Persistent Surveillance Systems" in Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010, edited by Jerome J. Braun, Proceedings of SPIE Vol. 7710, 2010.
3. Vinayak, E., Shirkhodaie, A., and Rababaah, A., “Context-Based Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems, 13th Int. Conf. on Information Fusion, 26-29 July 2010, EICC Edinburgh, UK.
4. Rababaah, H., Shirkhodaie, A., “Semantic Labeling of Non-Stationary Vehicular Acoustic Events in Persistent Surveillance Systems,” 13th Int. Conference on Information Fusion, 26-29, July 2010, EICC Edinburgh, UK.
5. MURI First-Year Research Progress Review Meeting, University of Buffalo,
6. MURI First-Year Research Progress Review Meeting, University of Buffalo, Buffalo, NY, September 27, 2010.

2011 Presentations:

1. Vinayak Elangovan, Haroun Rababaah, Amir Shirkhodaie, “Acoustic Semantic Labeling and Fusion of Human-Vehicle Interactions,” SPIE Defense, Security, and Sensing Conference, Signal Processing, Sensor Fusion and Target Recognition XX, April 25, 2011, Orlando, FL.

2. Vinayak Elangovan, Amir Shirkhodaie, "Context-based semantic labeling of human-vehicle interactions in persistent surveillance systems," SPIE Defense, Security, and Sensing Conference, Visual Information Processing XX, April 26, 2011, Orlando, FL.
3. Vinayak Elangovan, Amir Shirkhodaie, "A survey of imagery techniques for semantic labeling of human-vehicle interactions in persistent surveillance systems," SPIE Defense, Security, and Sensing Conference, Signal Processing, Sensor Fusion, and Target Recognition XX, April 25, 2011, Orlando, FL.
4. MURI Second-Year Research Progress Review Meeting, University of Buffalo, Buffalo, NY, September 27, 2011.

2012 Presentations:

1. Vinayak Elangovan, Amir Shirkhodaie, "Team Activity Analysis and Recognition Based on Kinect Depth Map and Optical Imagery Techniques*", SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-30, April 23, 2012, Baltimore, MD.*
2. Vinayak Elangovan, Amir Shirkhodaie, "Recognition of Human Activity Characteristics Based on State Transitions Modeling Technique*", SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-43, April 23, 2012, Baltimore, MD.*
3. Amjad Alkilani, and Amir Shirkhodaie, "A Survey on Acoustic Signature Recognition and Classification Techniques for Persistent Surveillance Systems*", SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-28, April 23, 2012, Baltimore, MD.*
4. Vinayak Elangovan, Amir Shirkhodaie, "Human Activity Discovery & Recognition based on State Transitions Modeling in Persistent Surveillance Systems," SPIE Electronic Imaging Science and Technology Conference, Multimedia Content Access: Algorithms and Systems VI, 22 - 26 January 2012, San Francisco, CA.
5. *Mohammad Habibi, and Amir Shirkhodaie, "A Survey of Visual Analytics for Knowledge Discovery and Content Analysis," SPIE 2012 Defense, Security, and Sensing Symposium, Conference:* Signal Processing, Sensor Fusion, and Target Recognition XXI, *Paper number 8392-27, April 23, 2012, Baltimore, MD.*
6. Vinayak Elangovan, Amir Shirkhodaie, "Adaptive Characterization, Tracking, and Semantic Labeling of Human-Vehicle Interactions via Multi-modality Data Fusion Model," SPIE Electronic Imaging Science and Technology Conference, Multimedia Content Access: Algorithms and Systems VI, 22 - 26 January 2012, San Francisco, CA.
7. Amir Shirkhodaie, "Perceptual Semantic Labeling of Human-Vehicle Interactions (HVI)," 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2012, Baltimore, MD.

8. Amir Shirkhodaie, “Semantic Labeling of Human-Vehicle Interactions Via Acoustic Events Characterization and Inference,” 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2012, Baltimore, MD.
9. MURI Third-Year Research Progress Review Meeting, University of Buffalo, Buffalo, NY, September 27, 2012.

2013 Presentations:

1. Shirkhodaie, A., Elangovan, V., Habibi, M. S., and Alkilani, A., “A Decision Support System for Fusion of Soft and Hard-sensor Information Based on Latent Semantic Analysis Technique”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
2. Elangovan, V., and Shirkhodaie, A., “A Robust Technique for Group Activities Recognition Based on Fusion of Extracted Features in Video Streams”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
3. Elangovan, V., Bashir, A., and Shirkhodaie, A., “A Multi-attribute Based Methodology for Vehicle Detection & Identification”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
4. Elangovan, V., Alkilani, A., and Shirkhodaie, A., “A Multi-Modality Attributes Unmasking Scheme for Group Activity Characterization and Data Fusion”, IEEE Intelligence and Security Informatics (ISI), Seattle, WA, June 2013.
5. Alkilani, A., and Shirkhodaie, A., "Acoustic Recognition of Human-Object Interactions in Persistent Surveillance Systems", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
6. Presented at the Army Research Laboratory, April 2, Adelphia, MD.
7. US Army Research Laboratory Research Day Meeting, ARL, April 2, 2013.
8. Habibi, M. S., and Shirkhodaie, A., “Mining Patterns in Persistent Surveillance Systems with Smart Query and Visual Analytics”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.
9. MURI Fourth-Year Research Progress Review Meeting, Penn State University, State College, PA, September 24, 2013.

2014 Presentations:

1. Alkilani, A., and Shirkhodaie, A., “Acoustic Events Semantic Detection, Classification, and Annotation for Persistent Surveillance Applications”, SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.
2. Elangovan, V., and Shirkhodaie, A., "Knowledge Discovery in Group Activities Through Sequential Observation Analysis", SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.
3. Habibi, M. S., and Shirkhodaie, A., "Multi-attributed Tagged Big Data Exploitation for Hidden Concepts Discovery", SPIE Defense, Security and Sensing Conference, Baltimore, Maryland, April 2014.

4. US Army Research Laboratory Research Day Meeting, ARL, June 19, 2014.

a. Peer-reviewed conference proceeding publications – N/A
b. Manuscripts - N/A
c. Books and Book Chapters – N/A

Honors and Awards

Titles of Patents disclosed during the reporting period – N/A

Patents awarded during the reporting period – N/A

Graduate Students

Ph.D. Students	**Per Cent Supported**
Haroun Rababaah[1] (Fall 2009-Fall-2010)	100%
Vinayak Elangovan[1] (Fall 2011 – Summer 2014)	100%
Amjad Alkilani[1] (Fall 2011 – Spring 2014)	100%
Moath Obeidat[1] (Spring 2014)	50%
Mohammad Habibi[1]	0%
Master Students	
Jerry Sweafford[1] (Summer 2012, Fall 2013)	25%
Bashir Alsaidi[2] (Spring 2012, Fall 2012)	25%
Biniyam Chaka (Spring 2010)	25%
Fatemeh Vaziriborog	25%
Vinod Bandaru[2] (Spring 2013, Fall 2013)	25%
Total Number of Graduate Students	4.75

[1] Ph.D.

[2] MS

Post Doctorates

Post-Doc Students	Per Cent Supported
Haroun Rababaah[1] (Fall 2010-Summer 2011)	100%
Amjad Alkilani[1] (Spring 2014)	50%
Total Number of Post-Doc Students	**1.5**

Faculty

Name	Per Cent Supported
Amir Shirkhodaie	10.0 %
Total Number of Faculty:	1

Undergraduate Students

Name	Per Cent Supported
Brandon Journey (Fall 2009)	25%
Shad Stud (Summer 2011)	25%
Jamal Hasan (Summer, 2011)	33.3 %
Anthony Baker (Summer,2012)	33.3 %
Adriann N. Wilson (Summer,2012)	33.3 %
Daniel Scobey (Summer, 2013)	33.3 %
Ramon Gonzalez (Summer and Fall, 2013)	50%
Mark Thelen (Fall, 2013)	25%
Diarra Fall (Fall, 2013)	25%
Ayeke Tegegne (Summer, 2014)	25%
Brent Warner (Summer, 2014)	25%

David Potter (Summer, 2014)	25%
Pedro Tavares (Summer, 2014)	50%
Total Number of Undergraduate Students:	**3.75**

Student Metrics

The number of post-graduates & PhDs funded during this period …	5
The number of under-graduates funded during this period …	13
The number of undergraduates funded who graduated during this period	10
The number of undergraduates who graduated during this period with a degree in science, mathematics, engineering, or technology fields …	10
The number of undergrads who graduated during this period and will continue to pursue a graduate or PhD degree in science, mathematics, engineering or technology fields …	6
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0	4
Number of graduating undergrads funded by DoD funded projects	10
The number of undergrads who graduated during this period and intend to work for the Department of Defense …	5
The number of undergraduates who graduated during this period and will receive scholarships or fellowships to further studies in science, mathematics, engineering or technology fields …	6

Masters Degrees Awarded

Name	Department	Thesis/paper title	Date
Total Number:			**5**

Ph.D.'s Awarded

Student Name	Graduation Date
Haroun Rababaah	Fall 2010
Amjad Alkilani	Spring 2014
Mohammad Habibi	Spring 2014
Vinayak Elangovan	Summer 2014
Total Number of Graduated Ph.D.'s	**4**

Technology Transfer

- Presented at four times at the MURI Annual Research Progress Review Meetings.
- Presented 22 papers at the SPIE Defense and Security Conference, 2010 through 2014.
- Presented two times at Army Research Laboratory, Adelphi, MD, 2013, and 2014.
- Presented at the 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2012, Baltimore, MD.
- Presented at the 3r^{d} Annual Human, Light Vehicle and Tunnel Detection Workshop, April 23-24, 2012, Mississippi University.
- Presented one paper at the IEEE Intelligence and Security Informatics (ISI), Seattle, WA, June 2013.

3.4.3 TSU Research Accomplishment in the Fiscal Years 2009-2014

Tennessee State University research objectives on this MURI project included:

(1) Development of suitable taxonomy and ontology for recognition of human-vehicle interactions (HVI), human-human interactions (HHI), and human-object interactions (HOI).

(2) Development of robust architectural framework with appropriate supportive computational models and techniques for multi-modality hard sensor fusion.

(3) Conduct human-in-the-loop experiments for characterization and discovery of suspicious social networks and group activities based on the capabilities of newly developed architectural framework for multi-modality sensor fusion.

(4) Develop a method for attribute-based characterization, and semantic annotation of sensors observed social networks and group activities;

(5) Test and validate the efficiency and effectiveness of newly developed multi-modality sensor fusion techniques and algorithms.

A summary of our earlier research accomplishment can be found in [J.7]-[J.24]. A more detailed technical description of our research accomplishments for this period is also available from our publications [J.1]-[J.6]. The following briefly presents a summary our research accomplishments for this period.

3.4.3.1 Understanding of Group Activities Taxonomy and Ontology

Recognition group activities taking place in urban environments requires understanding of taxonomy under which such group activities (GA's) are realized. Physical sensors discretely observe such activities and each frame of data offers certain useful information about the nature of group activity, but it does not explain it comprehensively. To facilitate a comprehensive explanation of a group activity, knowledge of underlying context is a necessity. By considering the taxonomy of GA's, a clear path toward the perception of the GA's can be established based on which ontological approaches for processing sensor data can be more readily implemented.

The explicit and expressive semantics of an application area's concepts, together with their relationships represented through logical formalisms and inference, constitute a knowledge representation known as Ontology. Ontologies allow automated processing of

data and information in a logical, well understood, and predictable way. Due to nature of multi-modality sensing, and the fact that each sensing modality provides a different level of understanding of events happening in the environment, it is imperative to consider separate ontology for identifying different sensing modality. For example consider ontology associated with Group Activities. A typical group activity may involve one or more persons, interacting with each other, some vehicles, and/or some objects. An ontology scheme for identification of social networks from observed group activities is illustrated in Figure 77.

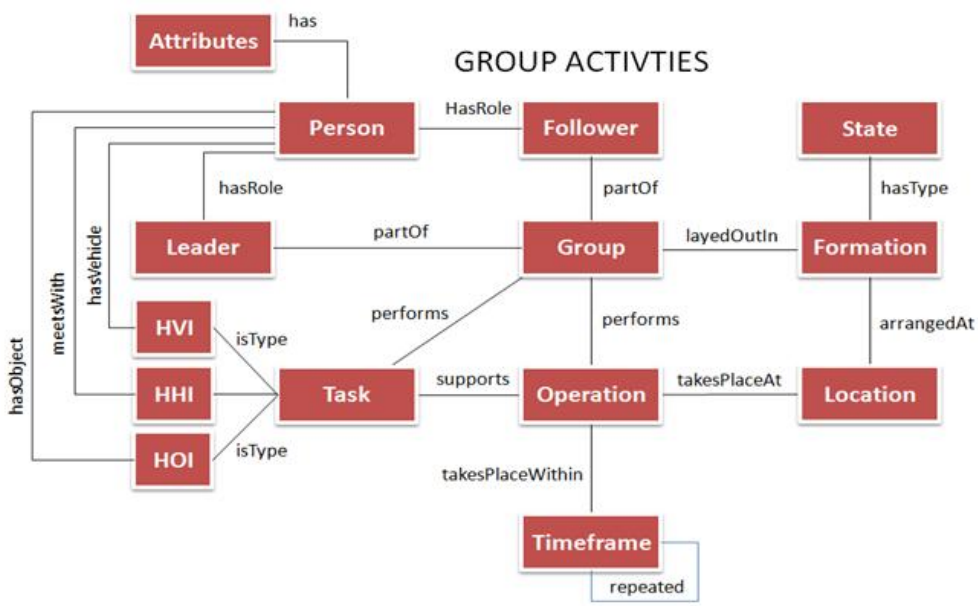


Figure 77: An Ontology Scheme for tracking Social Networking From Observed Group Activities

Figure 78 presents a model for matching an observed group activity to known ontological model with respect to taxonomy of operation and environmental factors of events detected by hard sensors. Figure 79 depicts how combination of different ontology can be implemented for spatiotemporal tracking of events and group activities.

Figure 80 presents a break-down the ontologies identifiable for two main groups of the physical (i.e., hard) sensors, namely optical imaging camera, and acoustic sensors. A typical group activity, at large, can be decomposed into three different types of interactions that we call them as: human-human interaction (HHI), human-vehicle interaction (HVI), and human-object interaction (HOI). For detection of each category of the interactions, certain attributes are required to be detected by the hard sensors. These attributes are associated with: Vehicle Detection, Human Detection, Object Detection, Group Pattern Detection, Group Activity Detection, Event Detection, Biometric Features Detection, Facial Features Detection, and Sound Detection. The attributes highlighted in yellow and green colors are detectable for image and signal processing techniques respectively. Those attributes that are highlighted in gray color represent categories of attributes that are unique and they hold one value, whereas those attributes that are highlighted in yellow or green colors are multi-value attributes and they are presented in Figure 81.

Attributes detectable for different sensor modalities are color-coded in Figure 81. Certain group activities attributes are suitable for detection for image processing techniques, and some others are more suitable for detection by signal processing techniques. In this period, we limited our scope of our hard sensor attribute detection to those highlighted for image and signal processing. The latter accommodation ensures a complete set of hard sensor attributes to be

achieved, though facial and biometric information are identified manually. The latter features, namely, facial and biometrics detection will be one scope of our research in the continuation of this present research effort. In the following discussion, we present the new techniques developed in this period for detection of group activity attributes based on cues detected by the hard sensors.

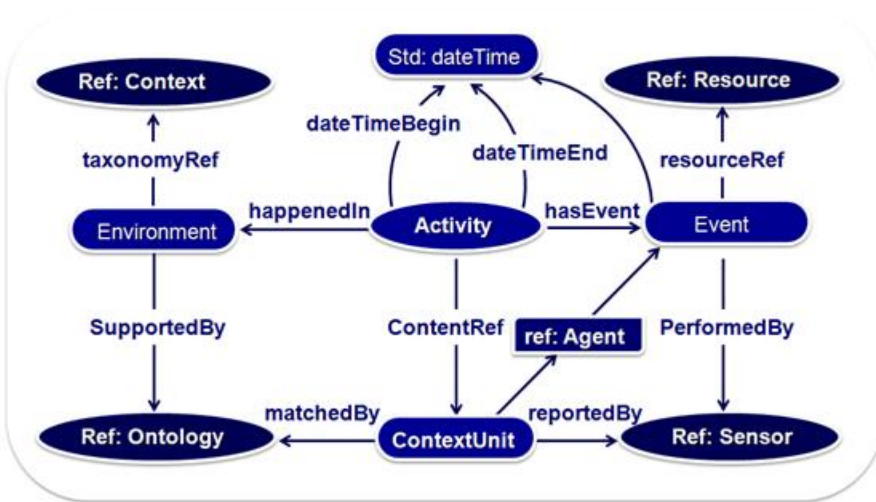


Figure 78: A Model for Matching an Observed Group Activity to Known Ontological Model With Respect to Taxonomy of Operation and Environmental Factors of Events Detected by Hard Sensors

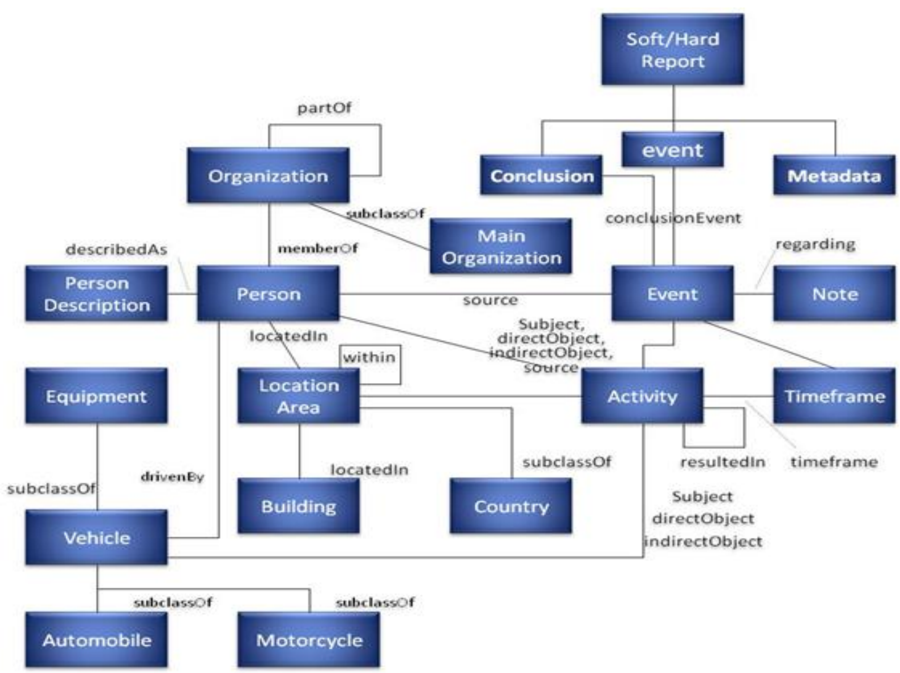


Figure 79: A combined ontology for spatiotemporal tracking of events and group activities

		ATTRIBUTES CATEGORIES					
Detection Operation ID	Hard Sensors Detection Description	Attribute-1	Attribute-2	Attribute-2	Attribute-4	Attribute-5	Attribute-6
1	VEHICLE DETECTED	VEHICLE TYPE	VEHICLE COLOR	VEHICLE SPEED	VEHICLE CURRENT STATE	VEHICLE PAST STATE	-
2	HUMAN DETECTED	CLOTH COLOR	HUMAN STATIC	HUMAN KINEMATIC	VEHICLE TYPE	OBJECT TYPE	OBJECT COLOR
3	OBJECT DETECTED	OBJECT TYPE	OBJECT SIZE	OBJECT COLOR	OBJECT SHAPE	OBJECT STATE	-
4	GROUP PATTERN DETECTED	PEOPLE COUNT	GROUP FORMATION	GROUP STATE	COUNT OF PEOPLE	COUNT OF PEOPLE	-
5	GROUP ACTIVITY DETECTED	GROUP ACTIVITY TYPE	PEOPLE COUNT	VEHICLE COUNT	OBJECT COUNT	VEHICLE TYPE	OBJECT TYPE
6	EVENT DETECTED	EVENT TYPE	EVENT FREQUENCY	EVENT PERIOD	EVENT SEVERITY	EVENT PROXIMITY	EVENT SEVERITY
7	BIOMETRICS FEATURES DETECTED	HEIGHT	GENDER	SKIN COLOR	HEAD COLOR	CLOTHING COLOR	
8	FACIAL FEATURES DETECTED	HEAD COVER	EYE COVER	MOUSTACHE	BEARD		
9	SOUND DETECTED	VEHICLE SOUND TYPE	HUMAN SOUND TYPE	OBJECT SOUND TYPE	ENVIRON. SOUND TYPE		

Figure 80: Categories of Attributes for Different Type of Hard Sensor Detection Capabilities.

Detection of prime events can be described by a combination of feature attributes. For example, a set of human activities (Inactive, Active, Walking, Running, and Sitting) can be differentiated using the combination of attributes of two or more features. As illustrated in the Figure 80, such features/attributes can take different types by which characteristics of events, entities, and nature of group activities are realized. Figure 81, on the other hands, demonstrates how an activity is comprehended upon detection of events characterizing a context-based activity. Note, the spatiotemporal representation of group activities are registered by space and time constraints.

ATTRIBUTES	TYPES OF ATTRIBUTES										
	1	2	3	4	5	6	7	8	9	10	11
VEHICLE TYPE	Sedan	Semi-truck	Truck	Van	Motor Bikes						
VEHICLE COLOR	Black	White	Red	Blue	Green	Orange	Yellow	Brown	Violet		
PEOPLE/OBJECT COUNT	One	Two	Three	Four	Five	Six	Seven	Eight	Others		
VEHICLE SPEED	Slow	Normal	Fast								
VEHICLE CURRENT STATE	Arriving	Parking	Departing	Engine On	Radio On						
VEHICLE PAST STATE	Arrived	Parked	Departed	Engine Off	Radio Off						
VEHICLE SOUND	Drived By	Arrived	Engine On	Engine Off	Departed	Door Open	Door Close	Hood Open	Hood Close	Trunk Open	Trunk Close
HUMAN SOUND	Talking	Yelling	Talking on Phone	Talking on Walkie-Talkie							
OBJECT SOUND	Radio Sound	Wooden Boxes	Cartoon Boxes	ToolBoxes	Glass Containers	Plastic Containers	Metallic Components	Metallic Trash Cans			
ENVIRONMENT SOUND	Wind Sound	Thunderstorm Sound	Rain Sound								
HUMAN STATIC POSTURE	Standing	Bending	Sitting	Laying Down							
HUMAN KINEMATIC	Motionless	Walking	Running	Jumping	Crawling	Pushing	Pulling				
CLOTHING/OBJECT COLOR	Black	White	Red	Blue	Green	Orange	Yellow	Brown	Violet		
OBJECT TYPE	Glass Container	Plastic Container	Metallic Object	Wooden Boxes	Cardboard Boxes	Light Weight	Medium Weight	Heavy Weight			
OBJECT SHAPE	Square Shaped	Long shaped	Round Shape	Unknown Shape							
OBJECT SIZE/EVENT PERIOD/MAG	Small	Med	Large								
OBJECT STATE	Person Dropped	Person Picked	Person Carrying	Taken From Vehicle	Placed in the Vehicle	left Behind					
GROUP ACTIVITY TYPE	Loading	Unloading	Object Exchange	Object-Dropping Unattended	Unattended Object Picking	Vehicle Exchange	Group Fleeing	Object Delivering	Vehicle Changing	Group Loitering	Vehicle Loitering
GROUP FORMATION	Merging	Scattered	United	Arriving	Joined	Marching					
GROUP STATE	Standing	Walking	Running	Jumping	Joined	Talking	Fighting	Screaming	Talking in Phone / Walkie Talkie		
EVENT TYPE	Intruding	Trespassing	Fighting	Fleeing	Shooting	Banging	Hiding				
EVENT SEVERITY	Low	Guarded	Warning	Elevated	Critical						
GENDER (BY APPEARANCE)	Male	Female									
SKIN COLOR	White	Light Brown	Moderate Brown	Dark Brown	Very Dark Brown						
HEAD COLOR	Black	White	Grey	Brown	Others						
HEAD COVER	Sports Cap	Military Cap	Worker Cap	Hat	Straw Hats	Others					
EYE COVER	Mask	Reading Glass	Sun Glass								
MOUSTACHE	Yes	No									
BEARD	Yes	No									

SENSING MODALITY COLOR CODES DESCRIPTION	
IMAGE PROCESSING	
ACOUSTICS PROCESSING	
IMAGE & ACOUSTICS	
MANUALLY PROCESSED	
HARD SENSOR EVENT SEVERITY PERCEPTION	

Figure 81: Types of Attributes Detectable by Different Hard (Physical) and Soft Sensors

3.4.3.2 Development of Techniques for Sensor Data and Information Fusion

Figure 82 presents TSU sensor fusion framework based on imagery and acoustic sensory data. Our proposed framework encompasses a number of research activities and supports the Hard Sensor (HS) data fusion aspects of this MURI research project. Particularly, for the objective of the project, we focused on development of algorithms and techniques that facilitate HS data processing and fusion via physics-based acoustic signals or observational imagery data. The framework supports task specific ontology and defines a generalized framework for generating feature vectors for detection, discrimination, and characterization of human behavioral activity pattern recognition as personal interaction either with other people, vehicles, or objects in the environment. The extracted feature vectors from different sensor modalities were used for teaching the system the human behavioral activities under different conceptual taxonomy. Furthermore, the proposed framework support decision support system for helpful for soft /hard decision fusion.

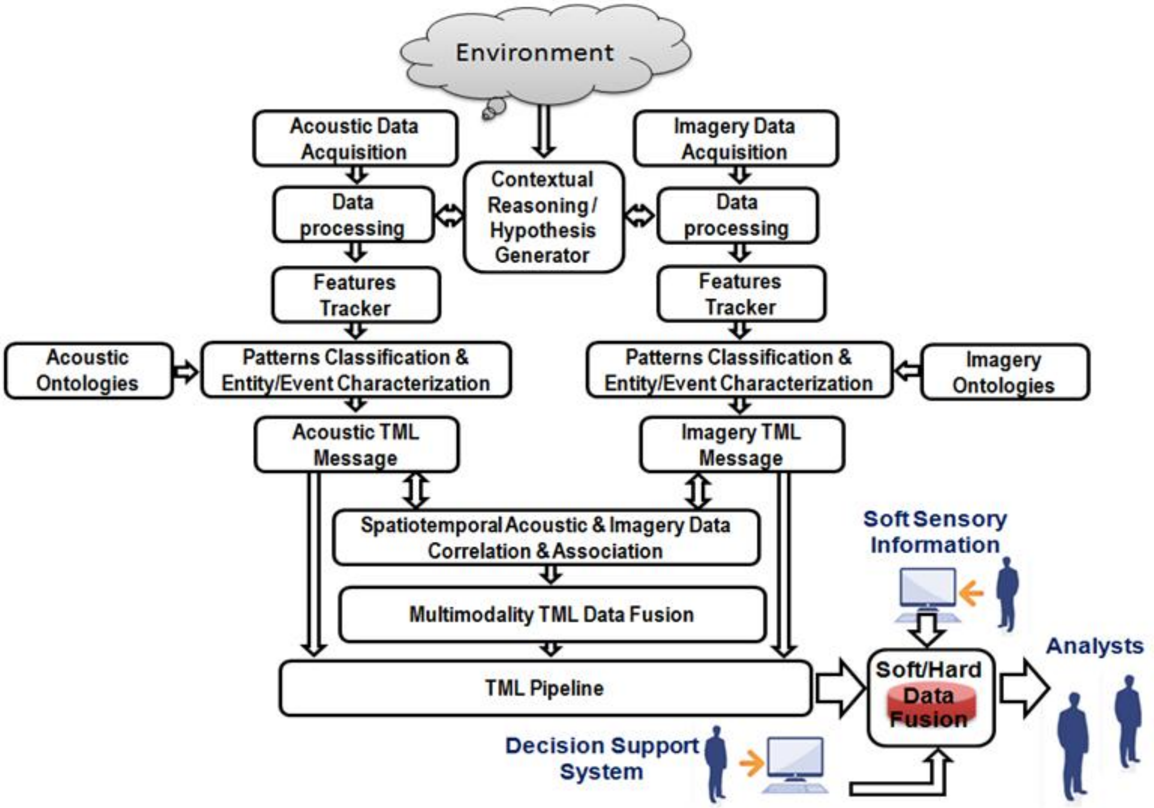


Figure 82: TSU Sensor Fusion Framework

3.4.3.3 HVI events

Human-Vehicle Interactions (HVI) refers to the type of activities that an individual may exhibit while using his/her vehicle (example: opening/closing vehicle doors, hood or trunk, turning on/off engine, arriving/departing at/from a vehicle parking location). Figure 83 shows some of the typical interactions a man is conducting with his vehicle.

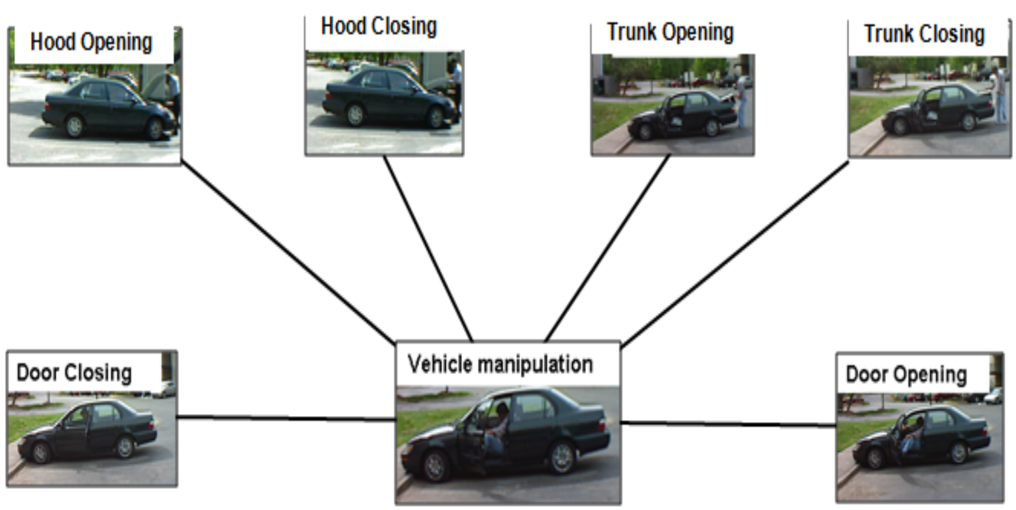


Figure 83: Sample of Human-Vehicle Interaction Events

In ZoV technique, the observed profile of vehicle of interest is divided into a matrix of connected cells, called zones. Each zone represents a key area of the vehicle for discriminating the HVI events. By analyzing the spatiotemporal HVI activities in such zones, one can ascertain type of potential/possible interactions that the person is involved with some degree of certainty. This spatiotemporal analysis is performed after detecting the orientation of the vehicle as discussed in previous section. As the vehicle arrives into a surveillance field of view and comes to a full stop, zoning is applied upon the vehicle profile. We partition the surrounding of vehicle into 20 different zones according to the spatial arrangement as illustrated Figure 84. Each zone denotes a specific location of the vehicle. For example: zone-15 belongs to the vehicle trunk region, zone-12 belongs to vehicle hood region, and etc. A volumetric CAD model had been developed for zoning the vehicle in different orientations as shown in Figure 85. The developed model can zone the vehicle in six different orientations namely, side view (front to back), side view (back to front), front view, back view, horizontal top view and vertical top view.

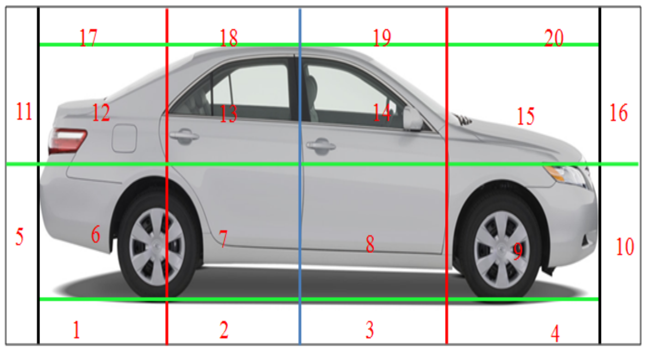


Figure 84: Zoning-of-Vehicle (ZoV) in Side View

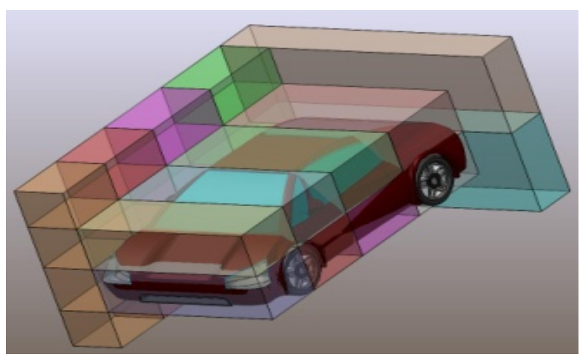


Figure 85: CAD model for ZoV in Different Orientations

Zoning-of-Vehicle helps in identifying 'whereabouts' of an event occurred around the vehicle. For example if a person opens the car hood, it can be identified as event of "Hood Open" occurred in "hood zone" and semantic messages are generated accordingly to describe the HVI [J.3]. Semantic labeling of events is generated with certain degree of confidence. To reduce the false alarm rate, information from two or more views of the vehicle (when available) may be fused and a probabilistic approach may be applied to further improve signal-to-noise ratio while reducing uncertainty associated with characterization of HVI events.

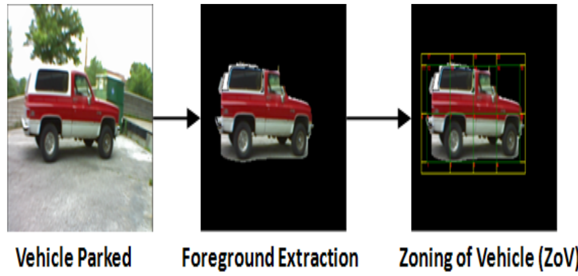


Figure 86: Zoning-of-Vehicle (ZoV)

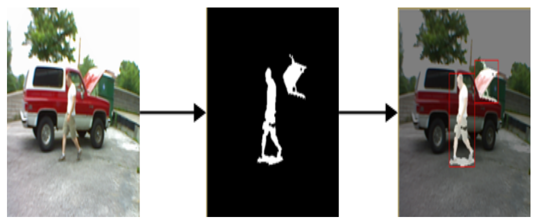


Figure 87: ZoV for Event Location Identification

Figure 86 shows the Zoning-of-Vehicle for a side view orientation of SUV and Figure 87 shows a sample of event location detection (i.e. person standing and vehicle hood open).

For analyzing a video stream of HVI activity, the events are detected and mapped to the developed HVI ontologies for predicting the series of actions involved in the scene. Many disciplines now develop standardized ontologies that domain experts can use to share and annotate information in their fields. By definition, the ontology is explicit formal specifications of the terms in the domain and relations among them. In other words, the ontology is involved with an iterative method of Knowledge-Engineering (KE) for a specific domain. The HVI Ontology development has several advantages including: (1) it facilitates common sharing of situational awareness, (2) it enables reuse of domain knowledge, (3) it makes domain assumptions explicit, (4) it separates domain knowledge from the operational knowledge, and (5) It facilitates to analyze domain knowledge. Figure 88 illustrates our hierarchical structure of the HVI ontologies. As demonstrated in Figure 89, our HVI ontology is presented in tree structure. The ontology tree is constructed based on clustering of ordered atomic events. One main advantage for presenting the HVI ontologies in the tree structure is that more complex ontologies can be developed based on simpler ontologies much more efficiently.

By focusing on metaphysics of HVI, we developed a rule-based system containing 200+ rules that links together what types of HVI are possible and what relations these events bear to one another to ensemble a situational awareness.

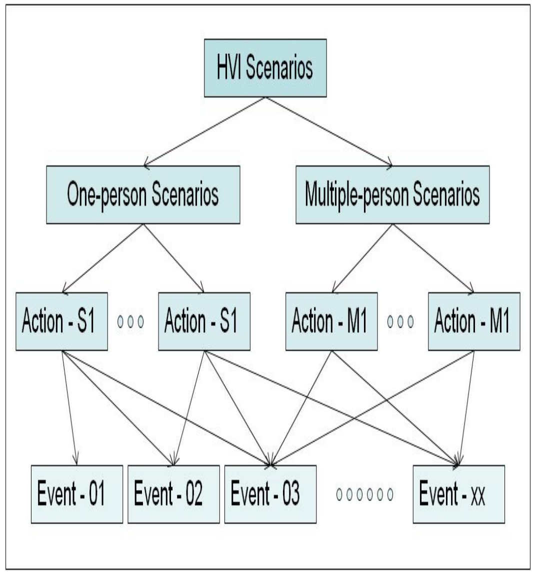


Figure 88: Illustrates the Hierarchy of HVI Ontologies

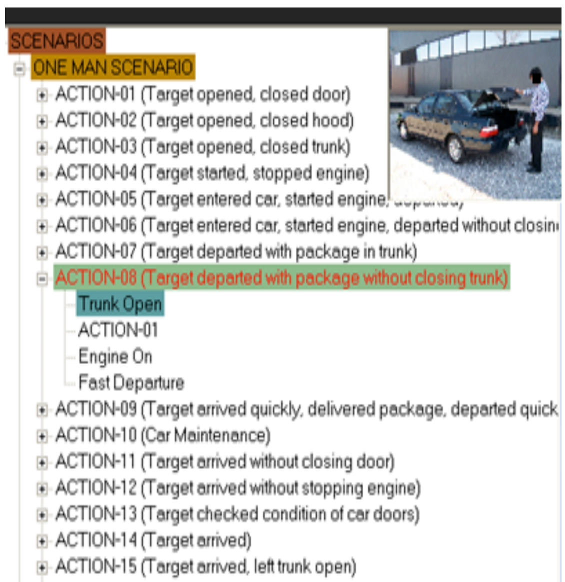


Figure 89: Complex HVI Ontologies

3.4.3.4 Human-Human Interactions (HHI) & Human-Object Interactions (HOI)

HHI is devoted to describing type of interactions a person may have with another person. Examples of HHI events are: shaking hands, hugging, waving hands etc. [J.8], [J.16]. To describe the characteristics of a human in a group activity effectively, essential attributes are considered, namely: Cloth color, Postures (example: standing, sitting etc.), Motion type (example: walking, running etc.), associated vehicle-ID, Interaction events with other humans (example: shaking hands, hugging etc.), and his social role in the scenario (example: Driver, Passenger, Subordinates etc.). HOI describes the type of interactions a person may have with an object. The techniques employed in detection of HVI events are also used in HHI and HOI event

detection with variations in feature parameters. Figure 90 shows the identification of human postures using the ZoV technique.

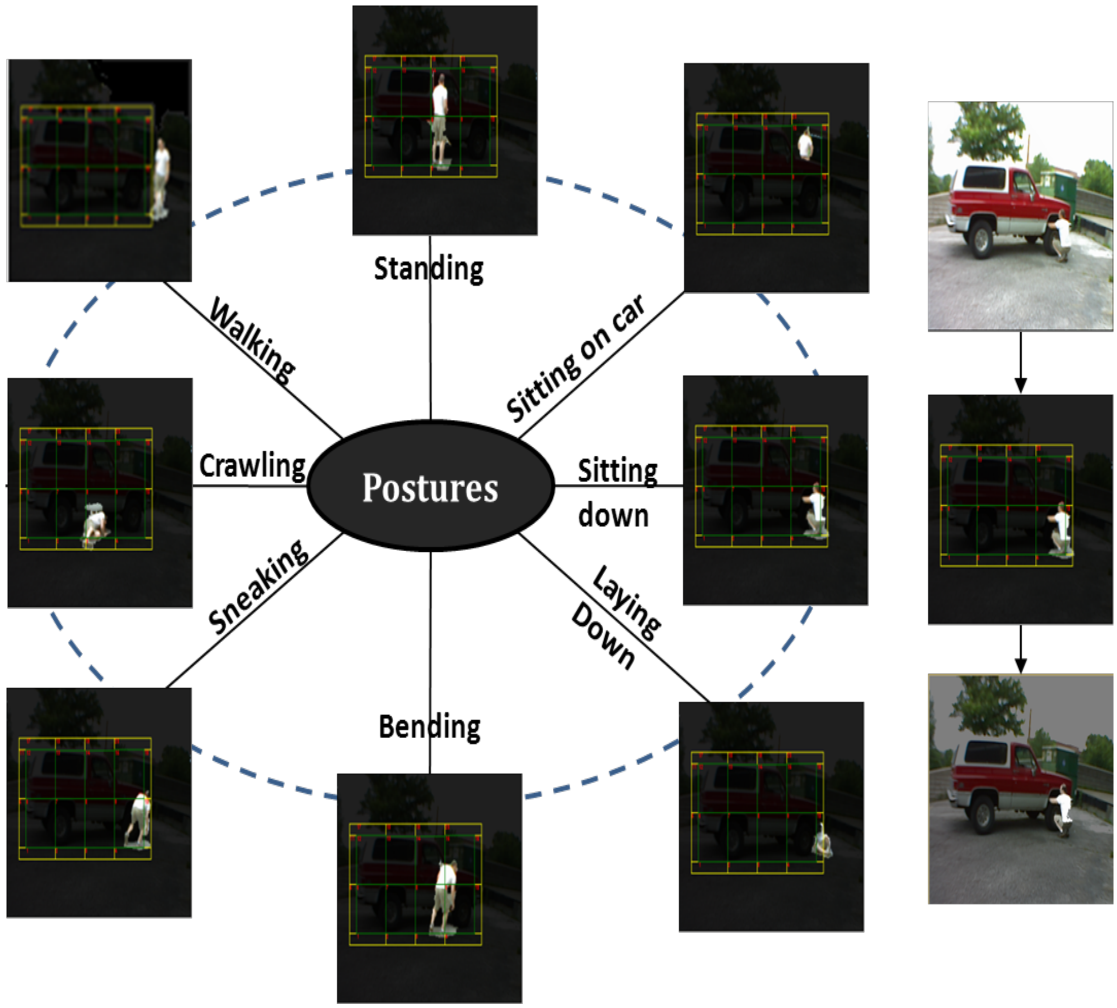


Figure 90: Posture Identification of a person around a vehicle using Zoning-of-Vehicle Technique

In the case of describing an object in a group activity effectively, the considered essential attributes are: Object color, Object Type (example: glass container, plastic container, metallic object, wooden box, card box etc.), Object Size (i.e. small, medium and large), Object Shape (i.e. square shaped, long shaped and unknown shape), events describing human interaction with objects (i.e. person dropped, person carrying, place in vehicle, taken from vehicle, person picked and left behind), and the corresponding ID of interacted human

Human-Human Interactions are recognized by isolating the detected targets and performing probe measurement technique for counting the heads for isolation of target entity. Local space correlation is performed on the isolated targets entity to detect the connectivity in order to determine the interaction. For example, shaking hands can be determined if there exists a blob connectivity between two individuals. After performing human isolation, the isolated images are matched with the collected templates.

Identification of human postures also enables in efficient detection of human-object interaction. For example, for an object removal, postures of human walking, standing, bending etc. are required. Human postures are detected by performing template matching. Training sets of human postures are collected from real scenarios and also geometric transformations like scaling, rotation etc., are applied on the collected samples for classification process [J.18].

In template matching, the correlation between two images is performed and the mathematical expression is given below:

$$\rho = \frac{COV(X_1, X_2)}{\sigma_1, \sigma_2} \qquad COV(X_1, X_2) = E(X_1 X_2) - \mu_{1,}\mu_{1,}$$

Where $\boldsymbol{\rho}$ is the correlation coefficient between two images: $\boldsymbol{X_1, X_2}$; and $\boldsymbol{\sigma_1, \sigma_2}$ are the standard deviations. The correlation factor ρ value varies between [-1,1]. If ρ is negative, then the two images are inversely correlated. If ρ is positive towards 1, then the images are strongly correlated. If ρ = 0, there is no correlation between the two images. The major steps followed in HHI and HOI processing:

For *each image*

→ *Identify Human TOI*

→ *Perform human posture classification - template matching or ZoV (if vehicle detected in same image)*

→ ***If*** *(number of Human TOI >1)*
Then {

- *Compute target-target directional vector and distance estimation*
- *Detect HHI events and generate semantic annotations*

}

→ ***If*** *(Object TOI = or > 1)*
Then {
For each object

- *Identify object profile through template matching and template differencing, and perform tagging*

}

→ ***If*** *(number of Human TOI >0 and number of Object TOI>0)*
Then *(Detect HOI events and generate semantic annotation)*

End

For detection of number targets, movements of individual and blob information is used. The profile of head is extraction by used top edge measurements on a binary image i.e. the perpendicular distance from the top edge of the image to the top edge of the binary image is calculated at various sequential points as shown in Figure 91. This feature vector is used to

detect the pattern of head assuming the usual shape of head is oval or circle. Using this feature, the detection of separation point of the individuals is also detected.

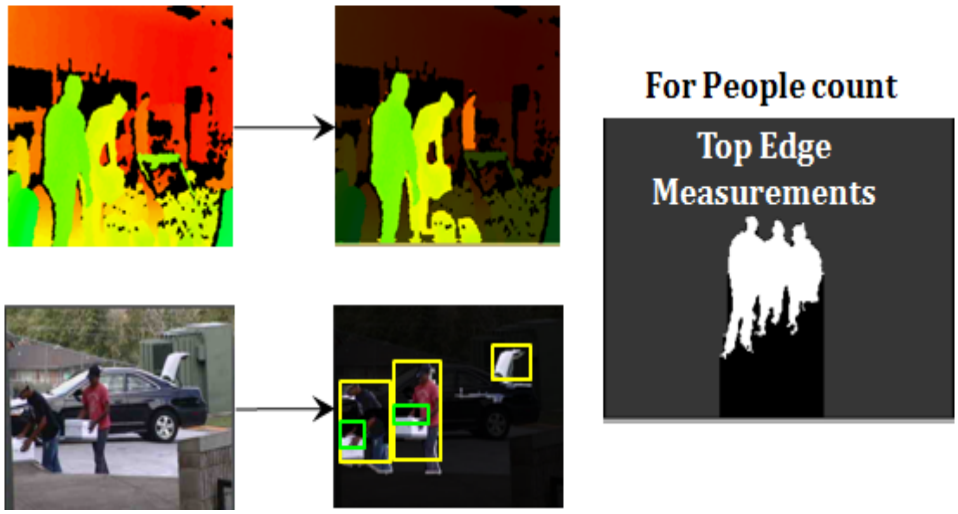


Figure 91: Samples of HHI and HOI Events Detection

3.4.3.5 Method of Target Detection

Detection and tracking of the target is efficiently done through the developed image processing techniques. The images to be processed are preprocessed (background removal) to refine the detection of the target. A sample of the target detection is shown below. The human blob is effectively detected through the background subtraction and the extracted foreground is also shown in Figure 92 and Figure 93 shows the detection of the various movements of the human target and the target blob can be tracked efficiently and displayed as shown in the figure. The detected BLOBs are processed through a pre-shape classifier to filter the Target of Interest (ToI). Pre-shape classifier classifies the blob into two different classes. Class-A holds the ToI i.e. vehicles and humans and class-B holds the noise i.e. unwanted blobs to be processed. Upon the detection of the blob, this classifier analyzes the ToI by extracting the relevant shape features such as blob area, blob elongation and circularity area to identify the target. Elongation and circularity ratios are used to give a sense of the nature of the shape. This Classifier uses the context information and metadata of the imagery sources as inputs for proper selection of the parameters. After a refinement of targets, the extracted target image in Class-A is fed to the developed Hamming Neural Network (HNN) for target classification as discussed in the following section. Figure 94 shows the messages generated for the target identification.

Figure 92: Foreground Extraction of Human Target

Figure 93: Foreground Extraction – Target Tracking

Figure 94: Screenshot of Target Identification Messages

3.4.3.6 Target Recognition and Classification

A scale invariant HNN had been developed for detection and classification of vehicles and humans. The developed HNN is invariant to the position, orientation and scaling of the target in a given image for classification. Image features of sedan, SUV, minivan and pickup truck are trained in the HNN. Training data contains 200 images of vehicle which includes 50 images of each class to be classified (i.e. sedan, SUV, minivan and pickup truck). The center of gravity i.e. centroid of the image is aligned with center of the image frame to generate a position invariant feature. The angle of moments of the image is used for constructing an orientation invariant feature. The scaling of the image is done by a measurement probe technique which measures the series of length between the horizontal edges from the centroid plane and mapping with the ratio of the measurement taken to the trained image's length and width parameters. Figure 95 shows the signature of the original vehicle images and the corresponding transformed image's signature. As seen in the figure, the signatures of the transformed images are close in similarity of the original images.

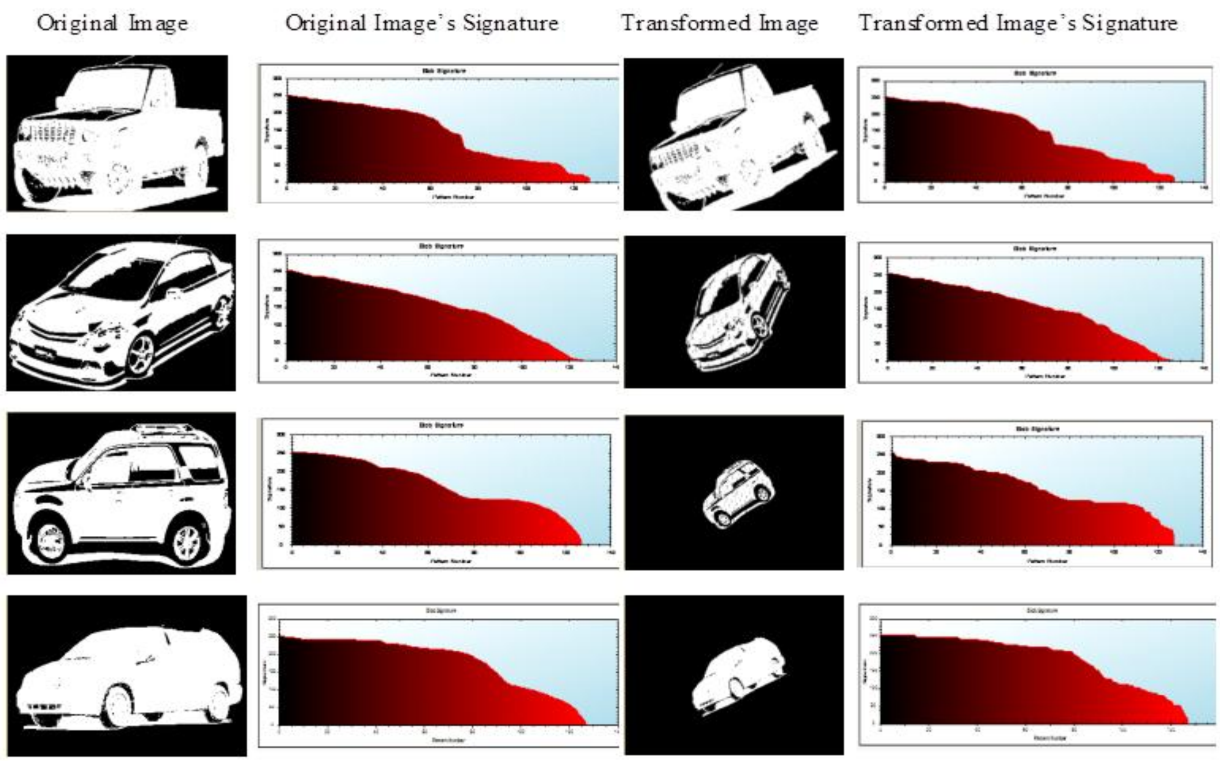


Figure 95: Signature of the invariant features of Vehicles

This invariant features signatures are used as inputs to the HNN for vehicle classification. For an efficient vehicle classification, the essential similarities to be identified between the target and the trained images are: (1) similarity in Hamming distance, (2) similarity in shape features and (3) similarity in the edge pattern as shown in Figure 96.

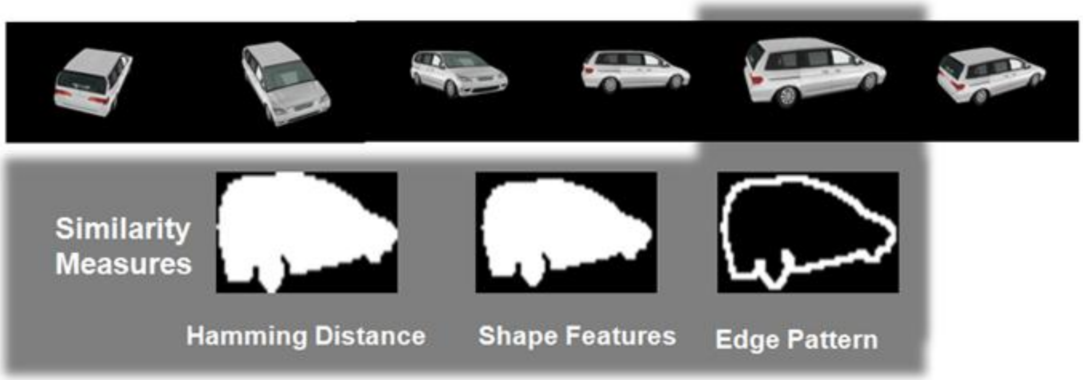


Figure 96: Processed Image of Minivan for Classification

In order to identify the similarities, we propose a Cascaded HNN classifier which cascades the multiple similarities features. A separate HNN classifier had been used for each similarity measure and a maximum likelihood is performed for identifying the type of the vehicle. Figure 97 shows the process of identification of vehicle attributes. Figure 98 also shows a sample of vehicles used as the training data. As seen in the figure, the HNN classifier identifies the four distinct types of vehicle.

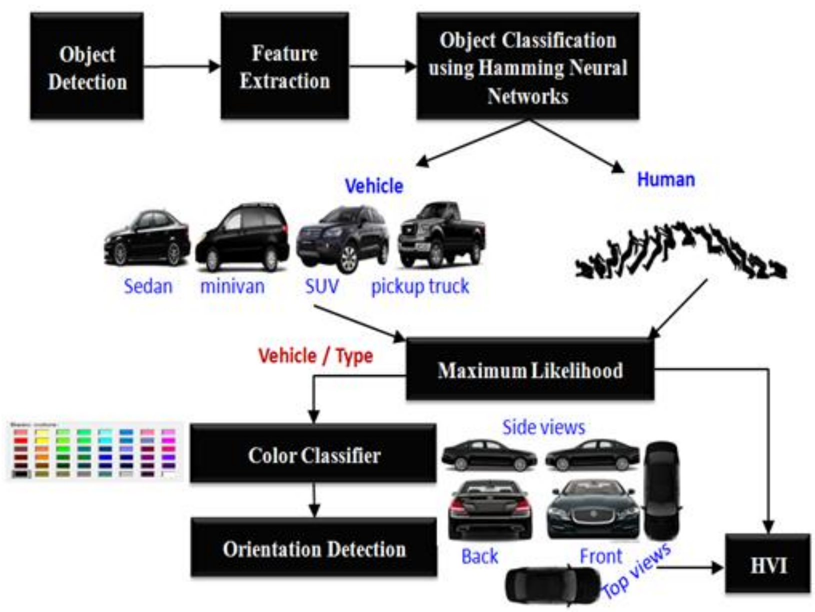


Figure 97: Process of Vehicle Attributes Detection

Figure 98: Sample of Vehicle Training Data

Figure 99 shows the screenshot of the application developed to Target Classification. The developed application for target classification can also detect and classify the color of the target by mapping to the known colors in the database. Three categories of color classifier are developed as shown in Figure 100. 'Solid' method detects the average color distribution in the entire target image. '2 segment' method divides the human target image into two segments namely upper body and lower body to identify the color of the shirt and pant respectively.

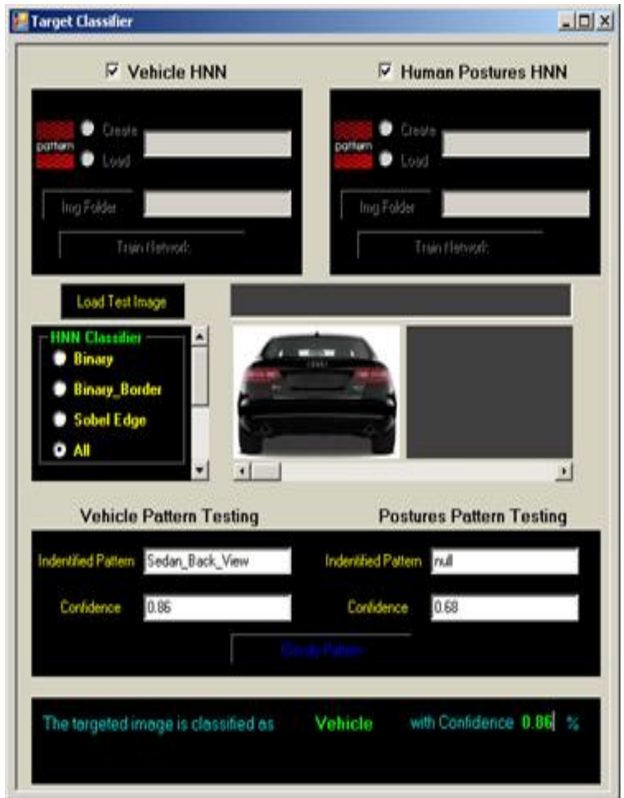


Figure 99: Target Classification Application Screenshot

Figure 100: Target Color Classifier

The color of the Vehicle is detected after the classification of type of the vehicle by the invariant HNN classifier. Ten most common vehicle colors have been used for color classification namely black, white, gray, green, blue, red, yellow, orange, pink and brown. In our work, we employ Region based color detection for detecting the vehicle color as shown in Figure 101.

Five provincial regions of (MxN) pixels are used from the body of the vehicle. The top and bottom region of the vehicle image are neglected since the color of the window / windshield and the tire region may possibly mismatch to the color of the vehicle. For each region, the mean of the RGB values of the total pixels (i.e. MxN) are computed as shown below:

$$\mu_R = \frac{\sum_{i=0,j=0}^{i=n,j=m} R_{ij}}{n \times m}, \quad \mu_G = \frac{\sum_{i=0,j=0}^{i=n,j=m} G_{ij}}{n \times m}, \quad \mu_B = \frac{\sum_{i=0,j=0}^{i=n,j=m} B_{ij}}{n \times m}$$

After finding the mean values in a region, the mean of RGB values from the five regions are computed if the RGB value of each region falls under the same Hue of the other regions i.e. the dominant Hue as shown below.

$$\mu_R = \frac{\sum_{i=0}^{i=r} R_i}{r}, \quad \mu_G = \frac{\sum_{i=0}^{i=r} G_i}{r}, \quad \mu_B = \frac{\sum_{i=0}^{i=r} B_i}{r}$$

where r is the number of regions selected with the dominant Hue.

The maximum likelihood of RGB value from the mean RGB values of the selected regions to the ten common colors is computed as the color of the vehicle.

Figure 102 shows the control options to be displayed/processed for Group Activity Recognition (GAR) and sample of vehicle and human detection.

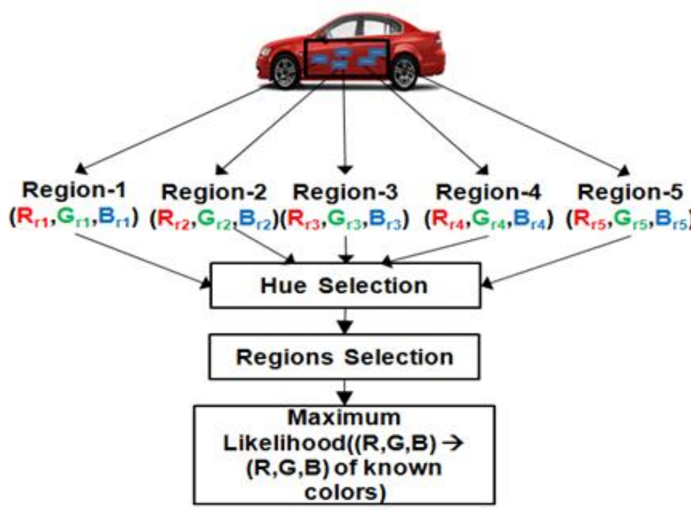


Figure 101: Vehicle Color Detection

The orientation of the vehicle is identified based on the detection of the vehicle color. The orientation of the vehicle is identified using top edge measurements and bottom edge measurements for obtaining the profile nature of the vehicle. The detection of the tire regions significantly helps in identifying the orientation of the vehicle. Bottom edge measurements are used in detecting the location of the tire region. Based on the vehicle color, three algorithms have been developed for detection of the tire region as shown in Figure 95, Figure 96, and Figure 97. Three distinct algorithms have been used since the color of the vehicle rims plays a vital role in algorithm selection. As demonstrated in the Figure 103, the color shade of the grey shaded vehicle matches with the color of the vehicle rims. Using the developed algorithms, it is seen that we are able to detect the pattern of the rims in black colored vehicle in the side view orientation.

Figure 102: Sample of Target Detection and screenshot of Group Activity Recognition (Gar) controls

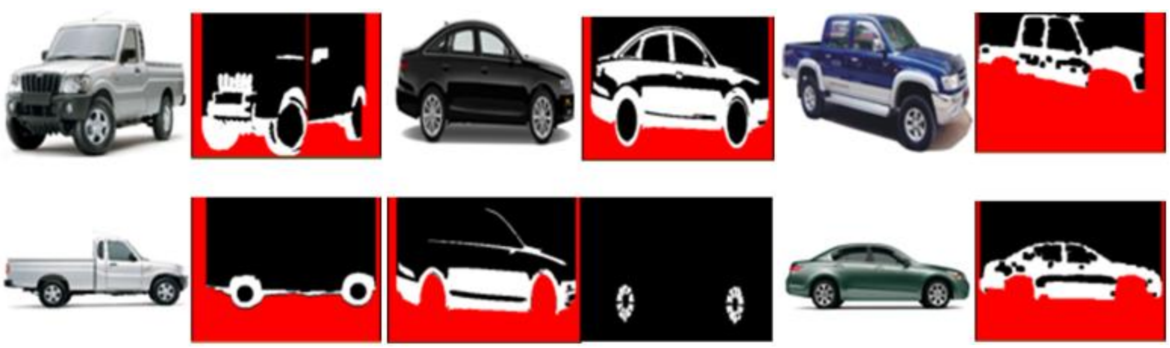


Figure 103: A- Grey Color Cars Orientation Detection. B- Black Color Cars Orientation Detection. C- Other Color Cars Orientation Detection

<u>Major Steps involved for Grey Car Orientation Detection:</u>

- *Remove RBG of the vehicle color*
- *Invert Image*
- *Remove small blobs*
- *Perform Bottom edge measurements*
- *Identify the Feature Vector and perform Maximum Likelihood on Feature Data set.*

<u>Major Steps involved for Black Car Orientation Detection:</u>

- *Perform RGB level-4 Segmentation*
- *Eliminate Grey shade*
- *Apply Default threshold*
- *Remove small blobs*
- *Perform Large closing*
- *Perform Bottom edge measurement*
- *Identify the Feature Vector and perform Maximum Likelihood on*

<u>Major Steps involved for Other Color Car Orientation Detection:</u>

- *Perform Mid Grey Threshold*
- *Remove small blobs*
- *Perform Large closing*
- *Perform Bottom edge measurement*
- *Identify the Feature Vector and perform Maximum Likelihood on Feature Data set.*

Feature Data set.

Upon the detection of the tire region, the distance between the center of the tires and to the distance between the edges of the vehicle body in the same plane is computed to determine the orientation i.e. side view or (front view or back view). Top edge measurements are used in identifying the top profile of the vehicle which in turn differentiates the side views orientation of the vehicle i.e. front to back side view or back to front side view. The detection of the exterior vehicle light dome are also used in differentiating the front view and the back view since most common vehicles have a unique color pattern of the light domes.

3.4.3.7 Human-Vehicle Interaction (HVI) Events:

For detecting the HVI events, we had proposed a Zoning of Vehicle (ZoV) technique, where each vehicle target is divided into different zones. Each zone represents the key area of the vehicle for discriminating the HVI events. By analyzing the spatiotemporal relationship between each zone and detecting human activities at such zones, one can ascertain type of potential/possible interactions that the human is involved with some degree of certainty. More details about the HVI can be found in our conference publications. The vehicle zoning helps in identifying 'whereabouts' of an event occurred around the vehicle. For example if a person opens the car hood, it can be identified as event of "Hood Open" occurred in hood zone and semantic messages are generated accordingly to describe the HVI [J.7]. Semantic labeling of events is generated with certain degree of confidence. To reduce false alarm rate, we fuse information from two or more views of the vehicle (when available) and apply a probabilistic approach to further improve signal to noise content and reduce uncertainty associated with characterization of HVI events. Figure 104 shows the application of ZoV technique for Back Side View of the Target Vehicle.

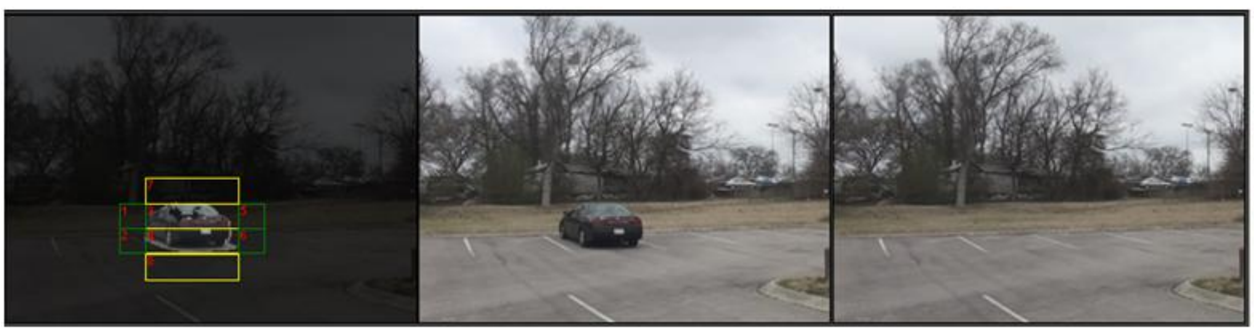


Figure 104: Zoning of Vehicle (ZoV) for Back Side View of the Target Vehicle

After the vehicle zoning had been applied, the whereabouts of the human target can be efficiently extracted as shown in Figure 105.

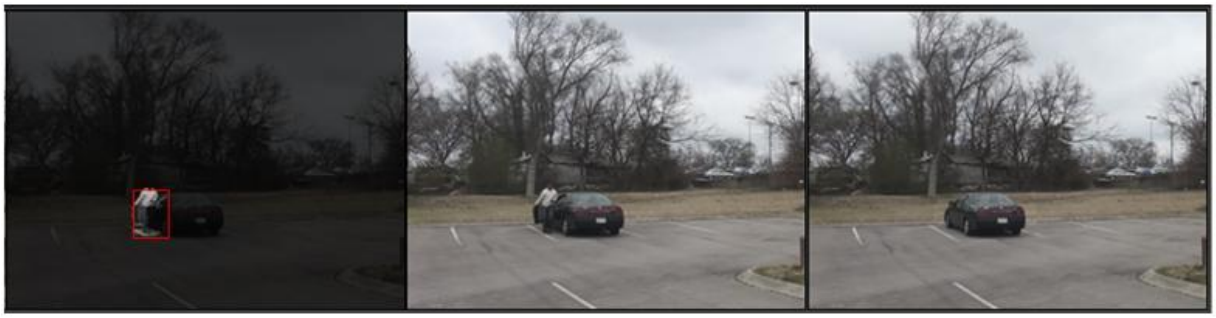


Figure 105: Human Exiting from the Vehicle

Figure 106 shows the detection area of the human exiting from the vehicle. By associating and correlating the detection of the blobs in the respective zone, the recognition of "Human exiting from the Front Door" can be concluded.

Figure 106: Human Exiting from the Vehicle through (ZoV) technique

Figure 107 shows the thumbnails extracted for the detected targets with the corresponding ID tags semantically annotated.

Figure 107: Thumbnails of Detected Targets

Figure 108 shows a sample of detection of the group interaction event. It is also noted the detection of another vehicle arrived in the scene. In order to detect the HVI events of the second vehicle, a proper background should be extracted. An Adaptive Foreground detection technique had been used in identifying the background as shown in Figure 109.

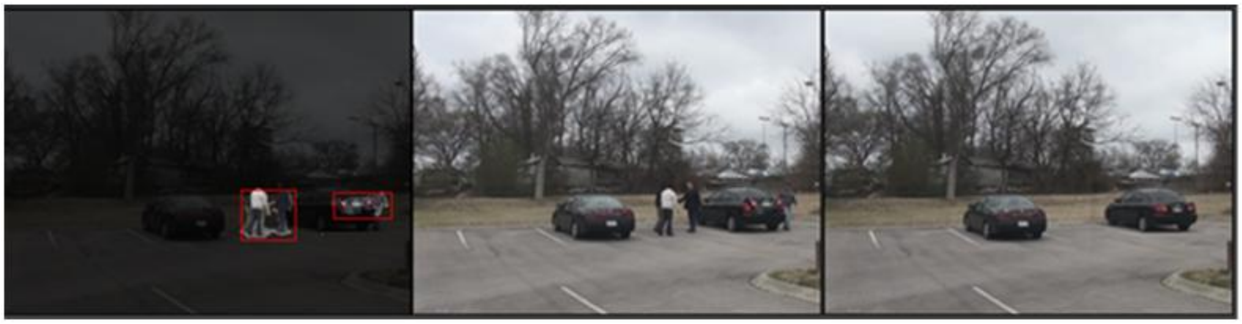


Figure 108: Detection of Group Interaction Events

Figure 109 shows the ZoV for the second vehicle in the side view orientation. It can be noted that the developed application can accommodate the zoning of multiple vehicles for detection of group activities involving multi-vehicles.

Figure 109: ZoV for Side View of Vehicle Target

3.4.4 Acoustic Signal Processing Techniques and Fusion

In our research, acoustic sensors were employed for recognition of sounds of things based on physical interaction of human with its environment. We experimented, primarily, with three types of sounds including: (1) sounds generated due to interaction of humans with environment, (e.g., sounds of walking with or without load, namely, sounds of human walking with or without carrying any heavy objects and sounds of human walking through different terrain conditions; (2) sounds generated due to interaction of human(s) with a vehicle (e.g., door opening/closing, trunk opening/closing, hood opening/closing, turning on/off engine; and (3) sounds generated due to interaction of human with the non-vehicular objects (e.g., sounds of lifting or dropping light/heavy objects, single/multiple objects, and metallic/non-metallic objects). Figure 110 presents the general taxonomy of acoustic sound we pursued for identification for Human Vehicle/Objects Interactions recognition and tracking. Figure 111 presents the taxonomy of HVI, HHI, and HOI considered for the scope of this project.

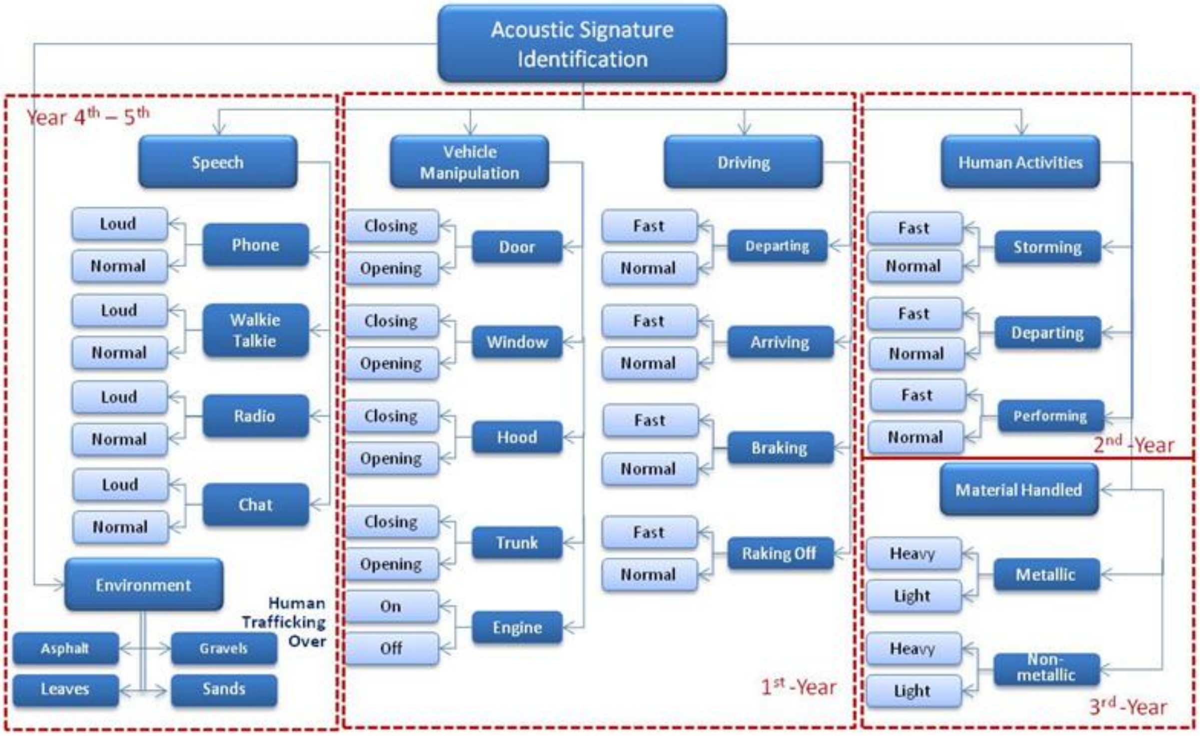


Figure 110: General Taxonomy of Acoustic Sounds Identification Based on Human Vehicle/Objects

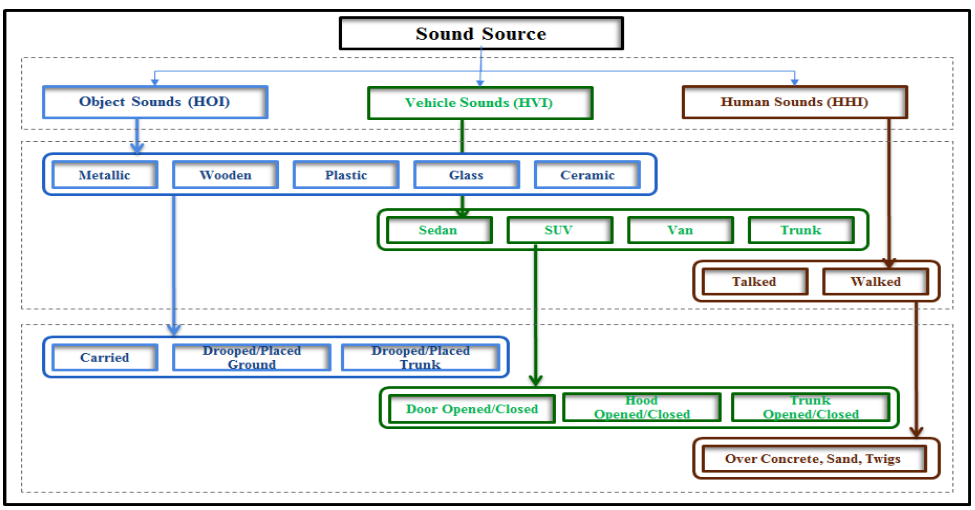


Figure 111: Taxonomy of Acoustic Sounds Considered For This Project

Interactions Recognition and Tracking

In this project, we were motivated to develop a technique for annotating Human-Object Interaction. In our earlier research work [J.5][J.7][J.10], we had presented a Neural Network approach for classification of human-vehicle interactions by training acoustic signal processors. More recently, we presented a survey of related signal processing applicable for human-object

interaction recognition [J.15]. In [J.18] we had also presented a technique for fusion of imaging and acoustic signals for detection of human-vehicle interactions.

In this period, we mainly focused on detection of metallic and non-metallic objects taken or removed from the environment. This problem was motivated since in many SYNCOIN operations, objects are being handled in containers that imaging sensors cannot identify their contents. However, acoustic sensors have potential to judge the nature of containers content based on the sound they make as the container box of such objects are being manipulated (i.e., lifted or dropped). Human ears can readily identify, for example, a box containing glass bottles from a box containing metallic parts when the box is dropped from a height. In this project, we were motivated to training a technique so that we can identify nature of content of a box based on the sound it generates. For the objective of this project, we developed a schema for acoustic sounds processing for Human Vehicle/Objects Interactions recognition, tracking, and semantic annotation. Figure 112 illustrates an overall perspective of processing stages of acoustic signals for detection, recognition, and tracking HVI, HHI, and HOI interactions. Figure 113 presents TSU's newly developed toolbox for Acoustic Signal Processing (ASP).

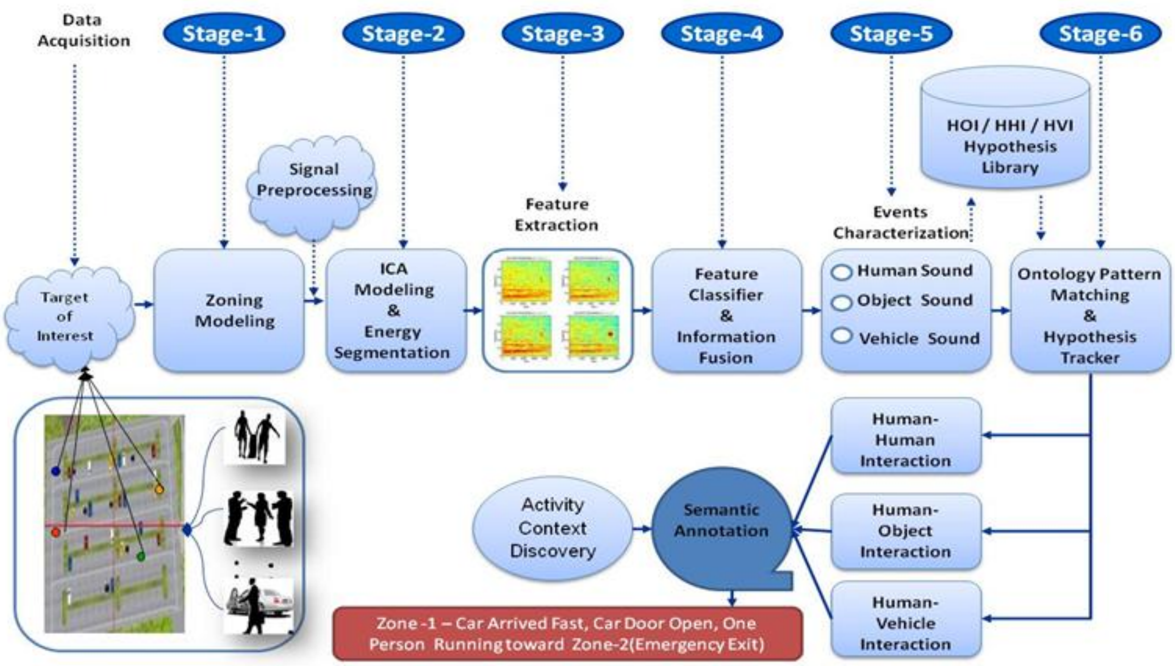


Figure 112: Stages of Acoustic Sounds Processing For Human Vehicle/Objects Interactions Recognition, Tracking, and Semantic Annotation

In this period, we conducted a series of experiments for detection of hidden objects in containers, boxes, and toolboxes. As illustrated in Figure 114, we considered a total of 252 acoustic signals from six different types of objects with three levels of weights and two levels of contents. All objects' sounds were collected by dropping objects from a height of one foot above the ground. Each experiment was repeated seven times for the purpose of training of acoustic signal processing techniques. Primarily, we chose six different types of objects including: 1) glass bottles of different sizes, 2) liquid containers (e.g., gas tanks, and water foundation

containers), 3) metallic pipes (e.g., from ½" to 4" diameter thin-walled aluminum pipes), 4) Wooden Boxes, 5) Carton Boxes, and 6) toolboxes (e.g., containing mechanical tools).

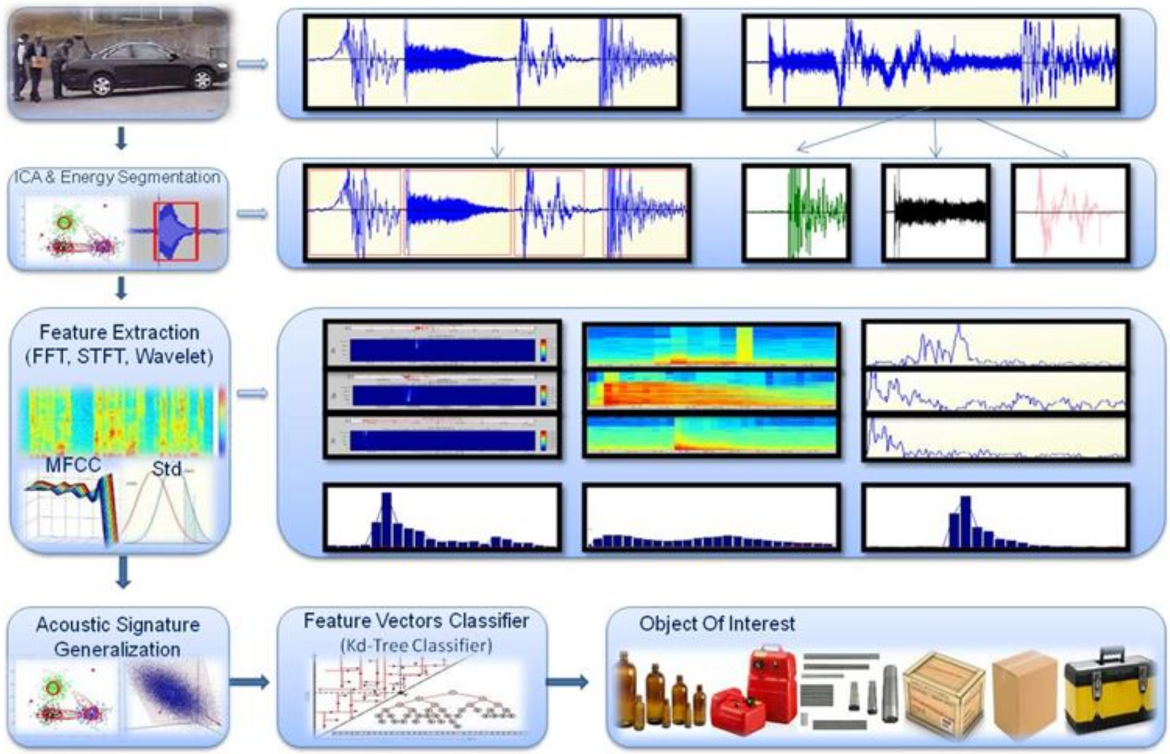


Figure 113: TSU's Newly Developed ToolBox for Acoustic Signal Processing

For each category of objects, experiments were conducted to generate sound waves as the object was being dropped from a height of one foot above the ground. The experiments were conducted on asphalted roads and sounds were recorded in open environment spaces with low-to-moderate ambient background noise. Initially, we attempted to classify the acoustic signals based on our earlier Kernel-based Spectral Similarity Matching techniques as illustrated in Figure 115. In this approach, after a pre-processing step of collected signals, a Gaussian Mixture Model was trained from characterization of spectral sound waves by clustering waveforms' principal component parameters of as detailed in [J.18].

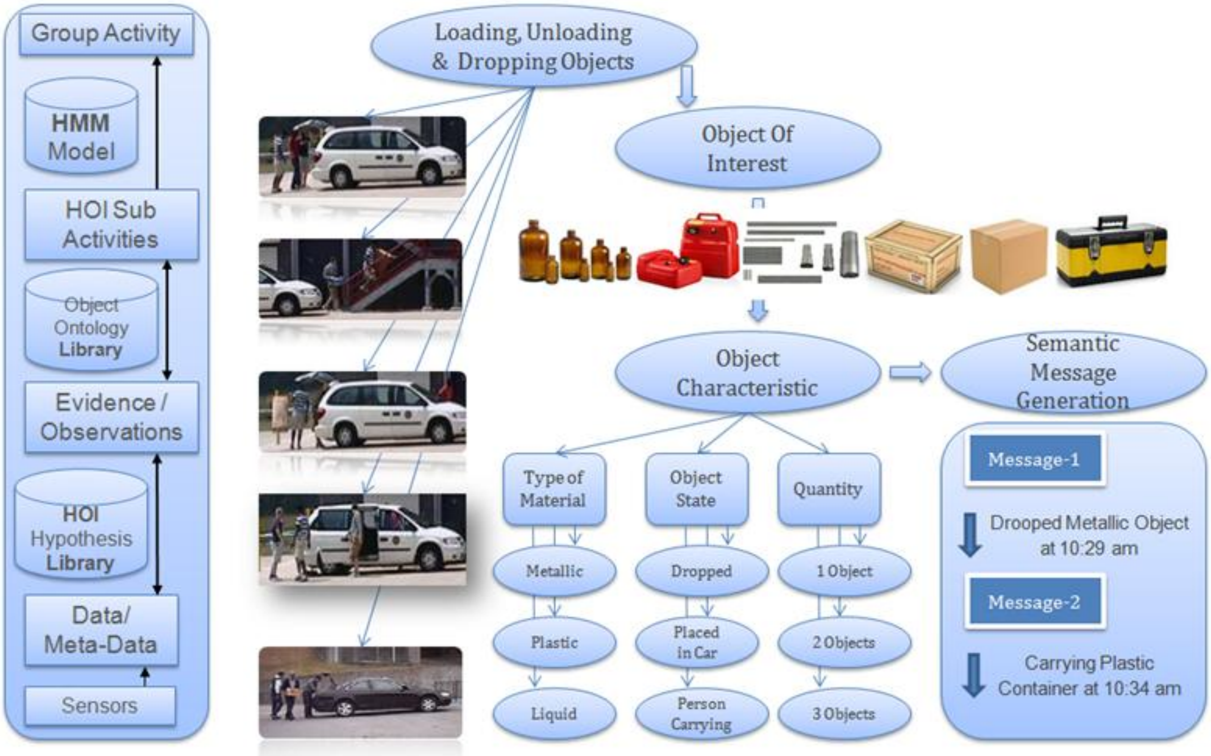


Figure 114: Details of Acoustic Signal Processing Experiments with Contented Objects

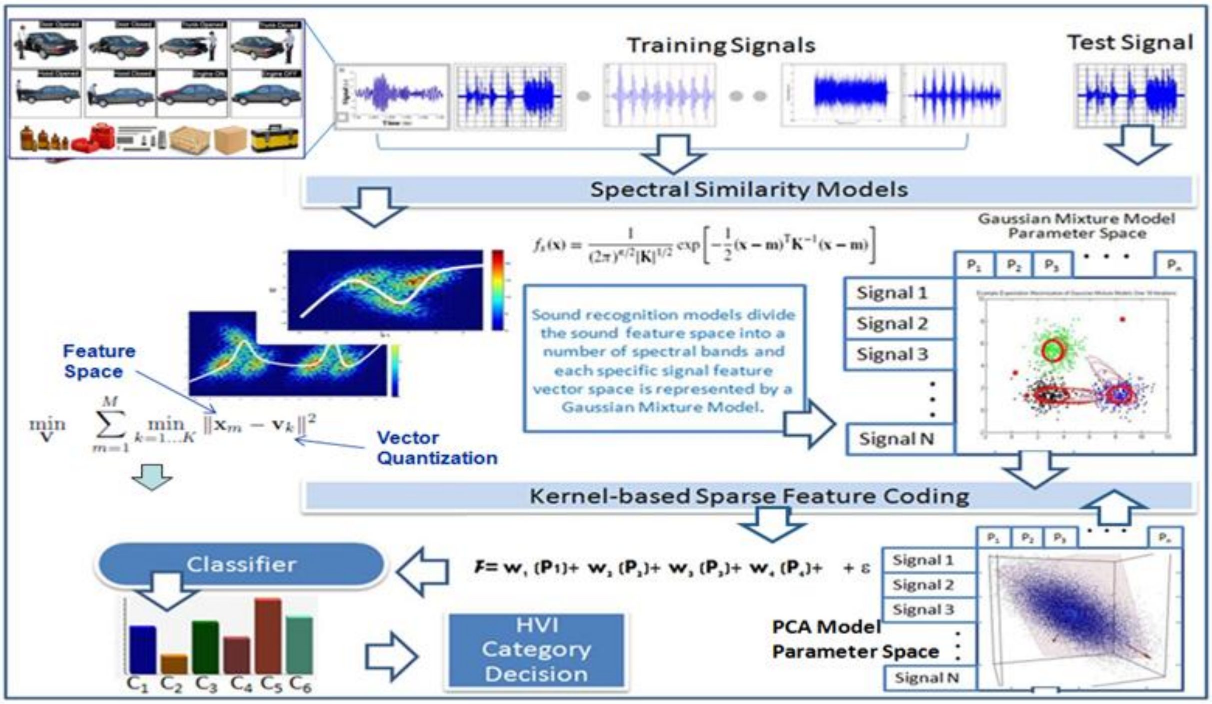


Figure 115: A Kernel-based Gaussian Mixture Model for Parameter Space Classification of Acoustic Signals

In another approach, we considered an ontology assisted technique for classification of Short-Time Fast Fourier (STFT) spectral of the sound waves. Based on this approach, initial the sound event is detected by monitoring signal energy level. Once signal energy level reaches a preset threshold, the signal is being monitoring till the energy strength of the signal drops below another preset threshold. Then, the spectral is bracketed into 32 bins and for each bin a FFT operator is applied to extraction of spectral parameters of the sub-signals. The acoustic sounds due to contented objects are partially non-stationary and partially stationary in nature. However, stationary harmonics are due to resonance frequencies of excited objects in the container. For examples, glass bottles, liquid containers, and metallic pipes when excited they become vibrant at their resonance frequency in a certain frequency bandwidth. However, other objects such as wooden boxes, carton boxes and toolboxes rather produce non-stationary sound waves when they get excited. We embark on this disparity to cluster frequency parameters associated with the test objects. Figure 116 presents a revised version of our earlier signal processing techniques, that uses an acoustic sound ontology for annotation of acoustic sound after the acoustic sound is recognized and classified.

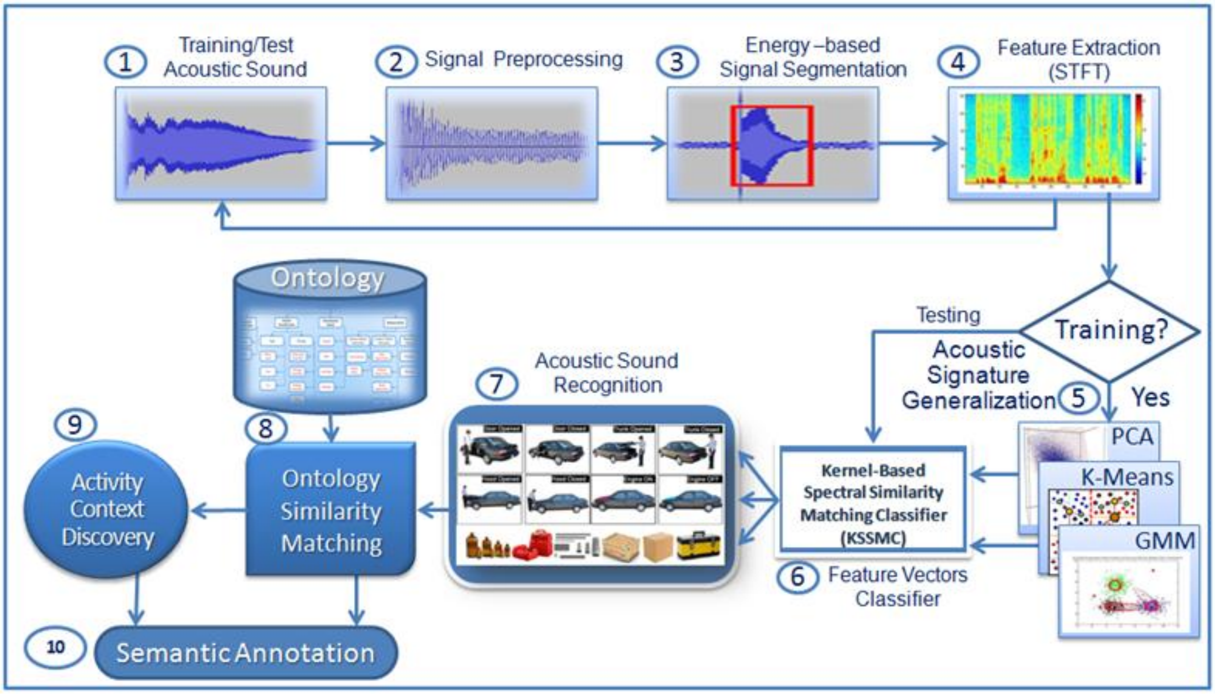


Figure 116: An Ontology Assisted Technique for Kernel-based Spectral Similarity Matching and Classification Based on PCA, K-Means, and Gaussian Mixture Model

One aspect of the acoustic work research this year was the continuation and completion of some work we had initiated on recognizing human interactions in Persistent Surveillance System (PSS). In the paper (Acoustic Signature Recognition Technique for Human-Object Interactions (HOI) in Persistent Surveillance Systems) which was published early in the year, we characterized different type of objects (Metallic, Glass, Wood, etc.) and their context (dropped on the ground or placed inside vehicle trunk) based on human manipulation of these objects. The objective of that paper was bi-folded. The first objective was to demonstrate how to achieve an improved situational awareness by means of acoustic sensors, particularly, in the area that objects of interest are hidden in containers that it complicates their delectability via surveillance cameras. The second objective was to demonstrate a technique for automatic semantic annotation of sound events.

3.4.4.1 Taxonomy of Human Interaction in PSS

Large order of acoustic signatures is generated due to operational activities of humans in the environment. In this research we divided such acoustic signatures into three major categories: Human-Vehicle Interactions (HVI), Human-Object Interactions (HOI), and Human-Human Interactions (HHI). Figure 117 presents a high level taxonomy of the three human interactions and how acoustic sensing modality can be effectively in detecting pertinent human activities.

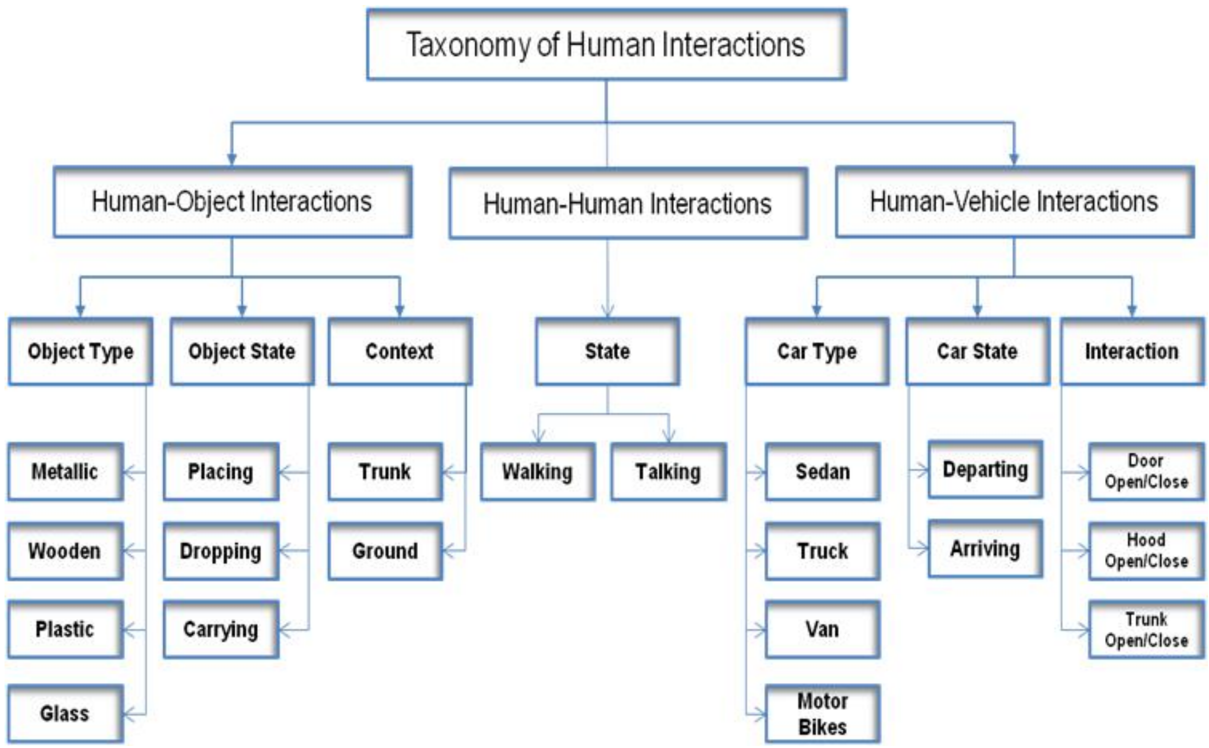


Figure 117: A High Level Taxonomy of Human Interactions

3.4.4.2 Background Noise Segmentation

The outdoor acoustic signals are typically contaminated with lots of background noises from the environment. Presence of noise typically compromises the effectiveness of the sound source detection and their classification reliability. Where the extent of signal-to-noise ratio is low, separation of two from each other becomes a difficult task. The recorded signal was smoothed using a low pass filter to remove high spatial frequency noise from a recorded signal.

3.4.4.3 Sound Source Segmentation

After suppression of noise from the sound source, an Energy-based segmentation was performed for isolating atomics acoustic events as shown in Figure 118.

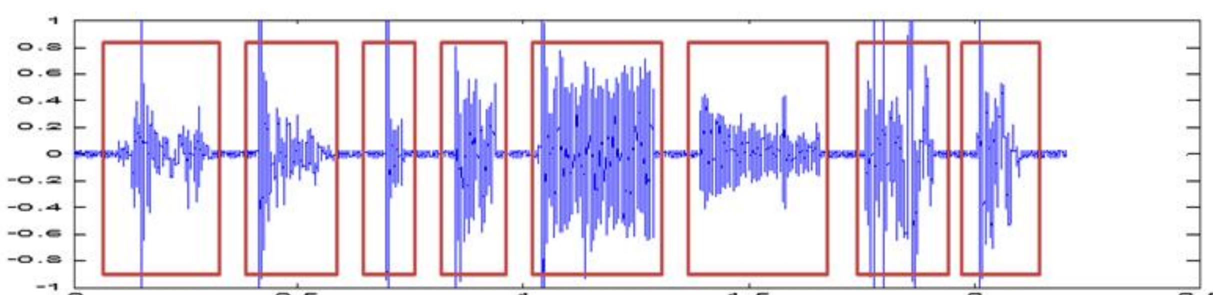


Figure 118: Isolated Atomics Acoustic Events

3.4.4.4 Training Dataset and Feature Extraction

For this research work, a rich variety of training sounds was obtained experimentally for different events of interest. Figure 119 shows different sounds of interest, and Figure 120 illustrates sample of experiments conducted and features extracted from different objects of interest. Feature Extraction entails various processing steps. In the first step, the salient signals extracted from the previous step are converted from their time domain to their frequency domain using conventional Fast Fourier Transform (FFT). Then we model this pattern using a frequency spectral envelope. The generated frequency envelopes further smoothed to reduce its dimensionality that, in turn, improves learning of frequency patterns representing unique acoustic signatures as shown in the Figure 120.

Figure 119: Sounds of Interest

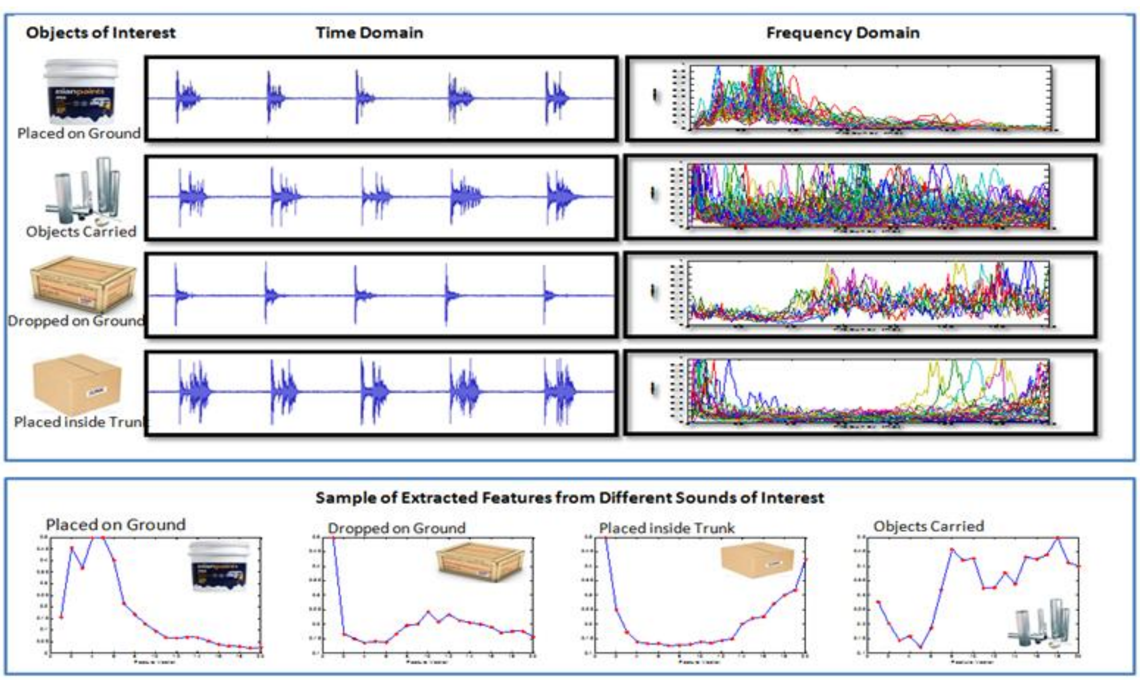


Figure 120: Sample of Experiments Conducted & Features Extraction

3.4.4.5 Features Extraction Interface

As part of this research work an acoustic toolbox has been developed using Matlab and Visual Studio C#, for acoustic events detection, classification evidencing the nature of human activities in the environment.

Figure 121 shows the feature extraction interface.

Training of sounds of interest can be done by uploading folder of sounds and selecting appropriate detection focus and type of event. Then the extracted features will be saved into a text file to be used for future classification of newly detected sounds. Figure 122 shows a text file for different sounds trained using the above interface.

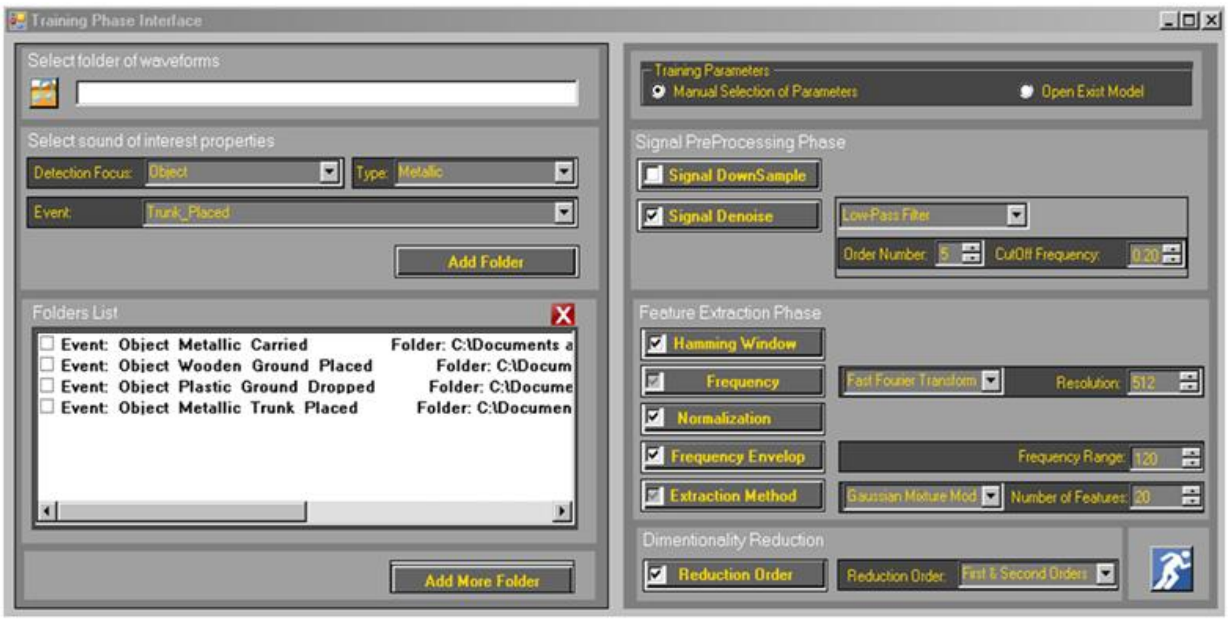
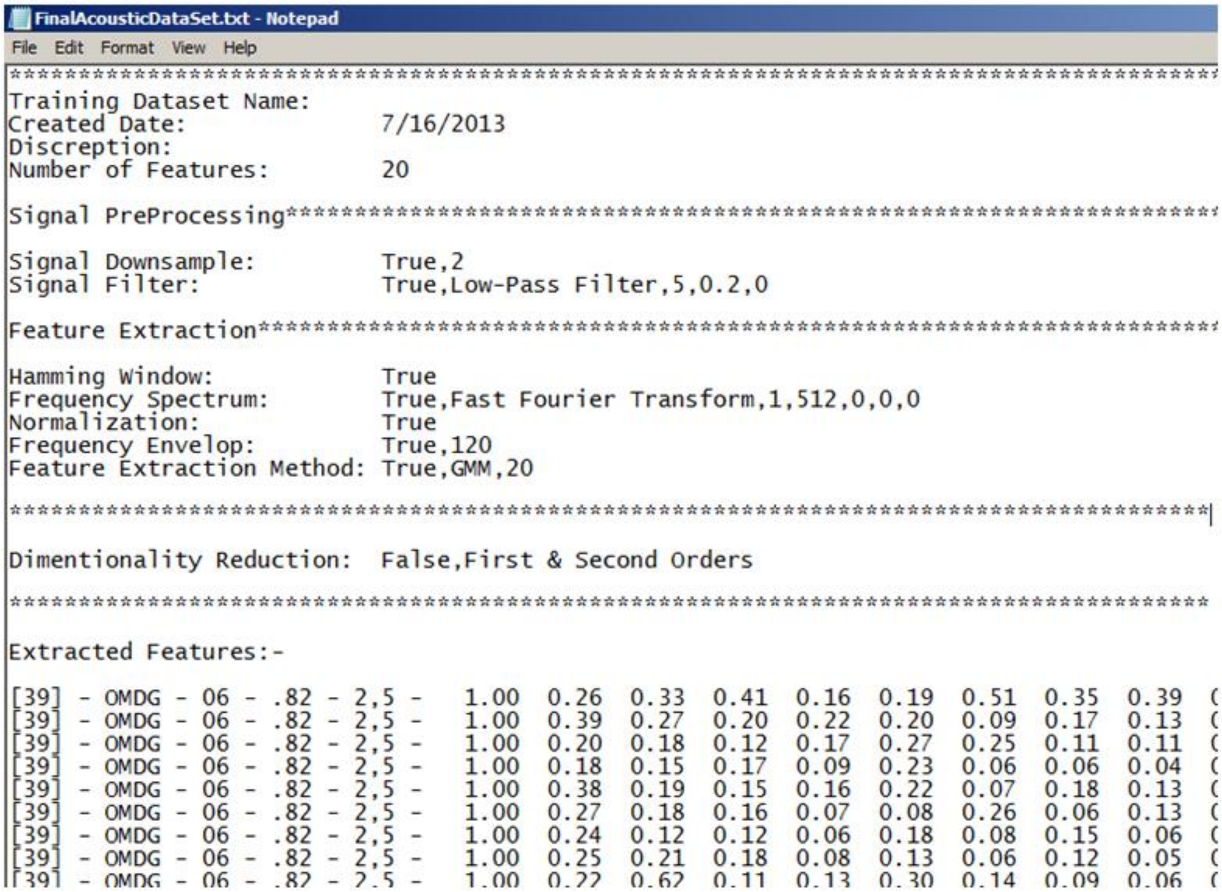


Figure 121: Features Extraction Interface



```
FinalAcousticDataSet.txt - Notepad
File  Edit  Format  View  Help
*****************************************************************************************
Training Dataset Name:
Created Date:            7/16/2013
Discreption:
Number of Features:      20

Signal PreProcessing*********************************************************************

Signal Downsample:       True,2
Signal Filter:           True,Low-Pass Filter,5,0.2,0

Feature Extraction***********************************************************************

Hamming Window:          True
Frequency Spectrum:      True,Fast Fourier Transform,1,512,0,0,0
Normalization:           True
Frequency Envelop:       True,120
Feature Extraction Method: True,GMM,20

*****************************************************************************************

Dimentionality Reduction:  False,First & Second Orders

*****************************************************************************************

Extracted Features:-

[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.26  0.33  0.41  0.16  0.19  0.51  0.35  0.39  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.39  0.27  0.20  0.22  0.20  0.09  0.17  0.13  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.20  0.18  0.12  0.17  0.27  0.25  0.11  0.11  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.18  0.15  0.17  0.09  0.23  0.06  0.06  0.04  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.38  0.19  0.15  0.16  0.22  0.07  0.18  0.13  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.27  0.18  0.16  0.07  0.08  0.26  0.06  0.13  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.24  0.12  0.12  0.06  0.18  0.08  0.15  0.06  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.25  0.21  0.18  0.08  0.13  0.06  0.12  0.05  (
[39] - OMDG - 06 - .82 - 2,5 -   1.00  0.22  0.62  0.11  0.13  0.30  0.14  0.09  0.06  (
```

Figure 122: Features Extracted of different sounds of Interest

3.4.4.6 Events Classification and Messages Generation

After the features been extracted, classification technique will be perform in order to identify to which of a set of categories a new detected signal belongs, on the basis of the trained sets of data. The classification interface developed is based on Correlation-based Template Matching (CB-TM) which relies on the statistical theory of correlation to find the best matching example that satisfies a threshold of confidence level. Figure 123 shows "Scenario Classification Interface".

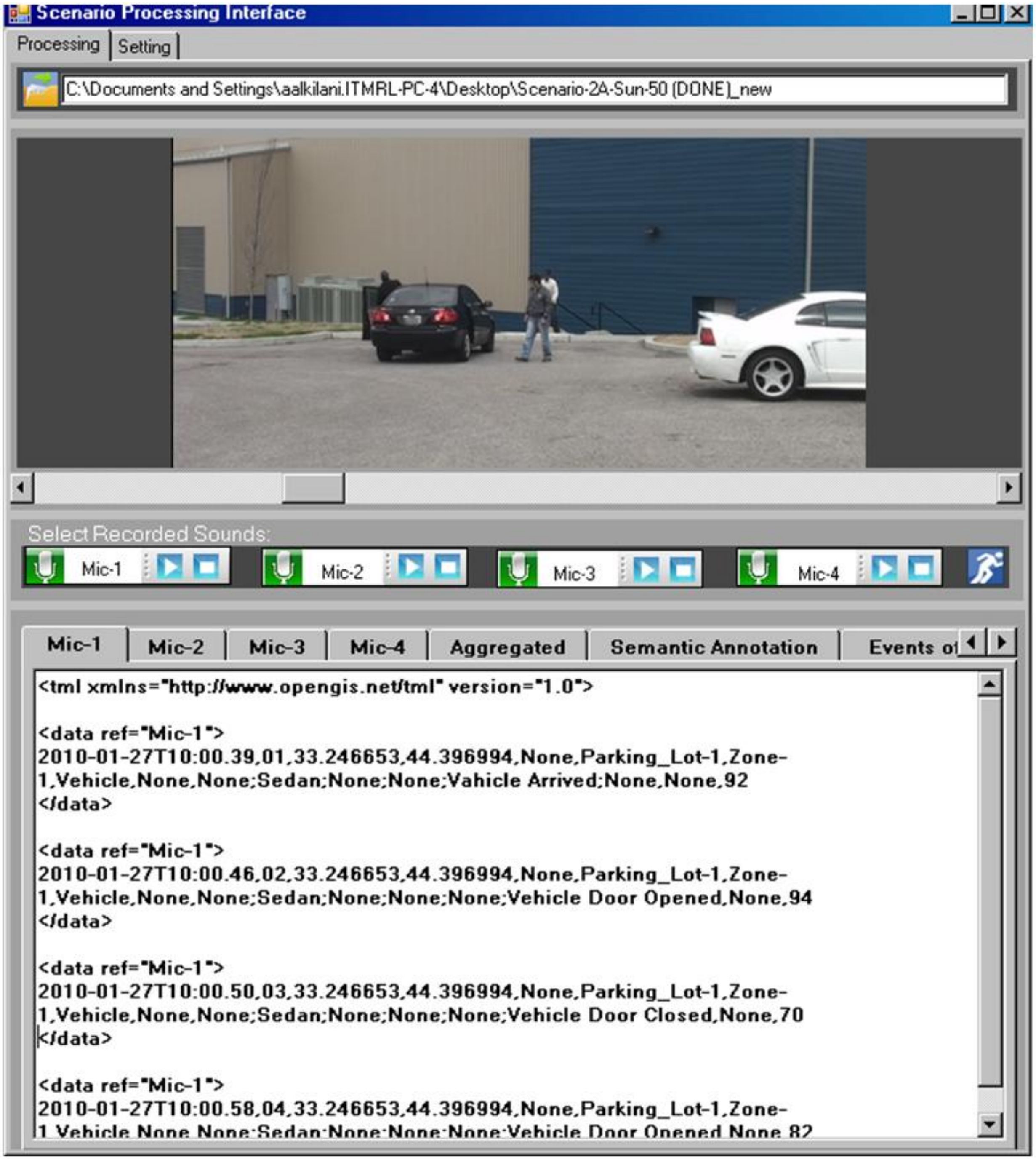


Figure 123: Scenario Classification Interface

The messages generated which describes the human activities is taking place, are semantically annotated using a Transducer Markup Language (TML) data format. Figure 124 illustrates a sample of the generated messages.

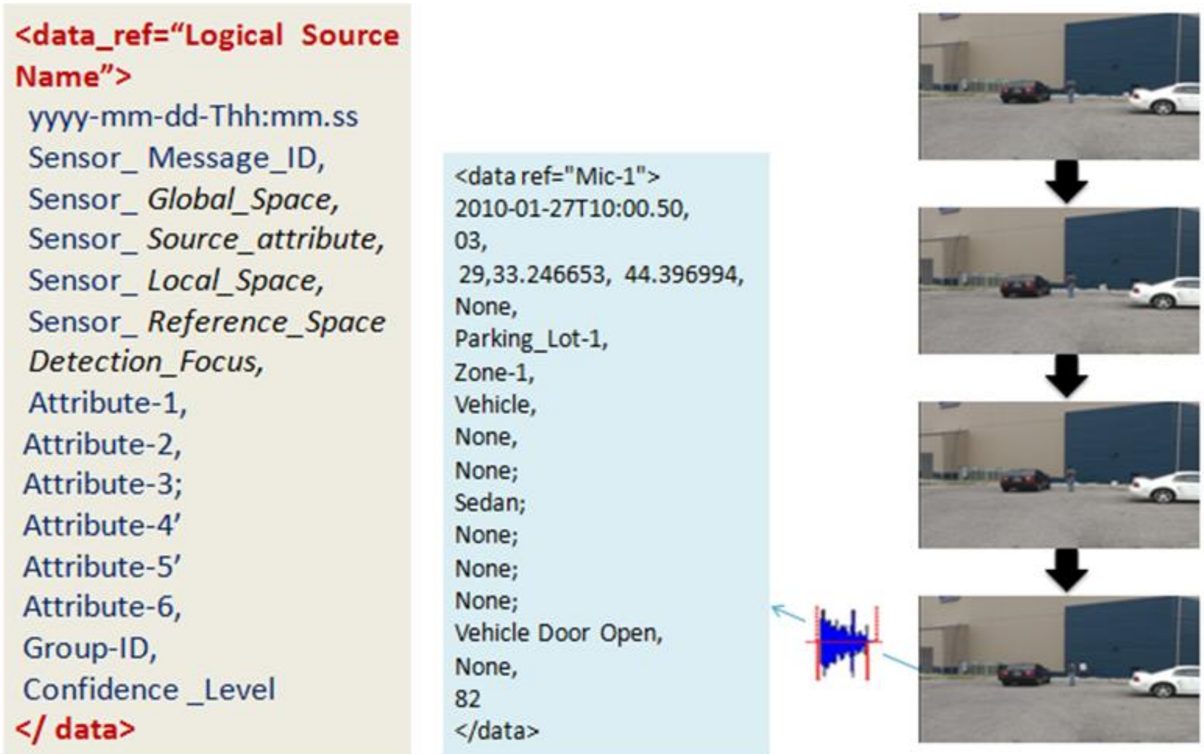


Figure 124: Semantic Annotation Format

3.4.4.7 Activity Recognition

Generally, outdoor activities are considered as a collection of interlinked salient events whose order of occurrences may represent a known ontology (i.e., a context). Using ontologies for activity recognition is a recent endeavor and has gained growing interest. Despite of many suggested paths, there is still a need for an explicit technique for defining outdoor activities. By streamlining activity definitions appropriate ontologies describing different order of activities can be established independent of underlying processing algorithms. This endeavor would ensure achievement of portability, interoperability, and reusability while allowing for sharing of both underlying technologies and systems. For example, by grouping such relational/sequential events, one can achieve an understanding of type of activity is taking place and there, the result of such a processing can become an effective activity prediction model. However, there is a need for mapping such observational events to specific concept that may be impending. Among many techniques for mapping a set of sequential observations (e.g., events) to certain outcome the Hidden Markov Models (HMMs) are the most commonly used methods in activity recognition. HMMs offer dynamic time warping, have clear Bayesian semantics, have well-understood training algorithms, and can model both large duration and small duration activities. Inherently, HMM represents a generative probabilistic model, which is a model that is, used for generating hidden states from observable data. In the AEDCA system, we used a trained version of HMM for Acoustic Activity Recognition (HMM-AAR). In HHM-ARR the hidden states (i.e., outputs) are known ontology states whereas the observations (i.e., inputs) are the detected atomic sound events. For computational aspect of the HMM-AAR developed, each event is originally

assigned a designated ID number according a scheme as illustrated in Table 32. Table 32A presents a list of unique atomic events with their assigned ID numbers. Ordered lists of some of these events represent a unique identifiable situation or simply revealing an acoustic ontological pattern. Table 32B presents the basic training ontological patterns by which the HMM-AAR is trained to recognize different patterns of sounds related to different activities pertaining to PSS.

Table 32: A- Unique Atomic Events (Sample). B- Ontological Patterns of Different Category of Activities

Event Name	Events ID
Vehicle_Door_Opened	1
Vehicle_Door_Closed	2
Vehicle_Trunk_Opened	3
Vehicle_Trunk_Closed	4
Vehicle_Hood_Opened	5
Vehicle_Hood_Closed	6
Vehicle_Parked	7
Vehicle_Engine_Started	8
Object_Trunk	9
Object_Ground	10
Object_Carried	11
Human_Walked	12
Human_Talked	13
Phone Rang	14
Walkie-Talkie Sound	15
. . .	. . .

Learning Patterns for HMM (Sample)

Sequential Events	Original Activity
14-13	Phone Conversation
13-14	Phone Conversation
13-15	Walkie-Talkie Conversation
15-13	Walkie-Talkie Conversation
7-1-2	Vehicle Arrived
1-2-8	Vehicle Departed
3-10-4	Unloading Objects
3-9-4	Loading Object
11-12	Moving Object
. . .	. . .

It is understood that recognition of sounds in much more complex environment, where many order of sub-activities may be taking place, introduces a much more challenging problem. Nonetheless, we recognize that some ordered sequences of sounds may have known interpretations. For instance, a person can conclude that a phone conversation is starting only if he/she is hearing a person talking a short while after hearing a phone ring. As appearing in this example, by analysis of spatiotemporal relations between the sequentially detected events, it would be possible to conclude a reasonable perception about the sound events heard. In order to insure a proper tracking of human activity using spatiotemporal information, each activity can be assigned a particular time-duration where in this time period, the associated events will be expected to occur. To support this tracking mechanism, a tokenizer activity handler has been developed. The task of this tokenizer is to automatically segment a sound stream into separated events while maintaining the temporal relationship of the sound events. In this process, time period between two adjacent events can be measured and use as a cue for correct identification of sound sources.

3.4.4.8 Experimental Results and Semantic Annotation

Two types of environments, indoor and outdoor, were considered for the collection of data and testing the AEDCA subsystems, each of which has its own unique attributes. In general, acquisition of acoustic data from the outdoor environment is inherently more challenging than collecting sounds in the indoor environments. There are a number of factors that contributes to this disparity. One such aspect is the inevitability of the wind blowing and dynamic ambient sounds (e.g., constant changing of traffic sounds), which increases the false alarm rate associated with acoustic sound classification and characterization. Also, the open outdoor space

significantly affects the quality and reliability of the collected sounds, particularly those measured from far distances. Moreover, the outdoor environment may result less multi-path concern that is a more eminent concern for the indoor situations. Most of the outdoor experiments were conducted at night time to minimize adverse traffic sounds. The developed framework was evaluated at two different levels (1) Event Classification: through the evaluation of the performance of the two classifiers, CB-TM classifier and k-d tree classifier, used by the engine of the acoustic semantic annotation. (2) Activity Recognition: through the evaluation of the performance of the ability of the HHM-AAR to track the occurrence of an individual activities taking place in the environment.

3.4.4.8.1 Performance Evaluation of Event Classifiers

In this section we compare performance of CB-TM and Kd-Tree classifiers. Table 33 shows a combination of HHI, HVI, and HOI considered for performance evaluation of the two classifiers.

Table 33: Acoustical Events Considered in the Database

Vehicle Events of Interest			**Object Events of Interest**			**Human Events of Interest**		
Detection -Focus	Type	HOI	Detection-Focus	Type	HVI	Detection-Focus	Type	Surface Type
Vehicle	Sedan	Door Opened/ Closed	Object	Glass	Placed/ Dropped Trunk	Human	Walked	Over Concrete
	SUV	Hood Opened/ Closed		Plastic	Placed/ Dropped Ground			Over Sand
	Van	trunk Opened/ Closed		Wooden	Carried			Over Twigs
	Truck			Metallic			Talked	
				Ceramic				

For each one of these sound of interest an array of related sounds collected experimentally and subjected to the classification process. Figure 125 illustrates a sample of different atomic sound events generated from closing of a sedan door.

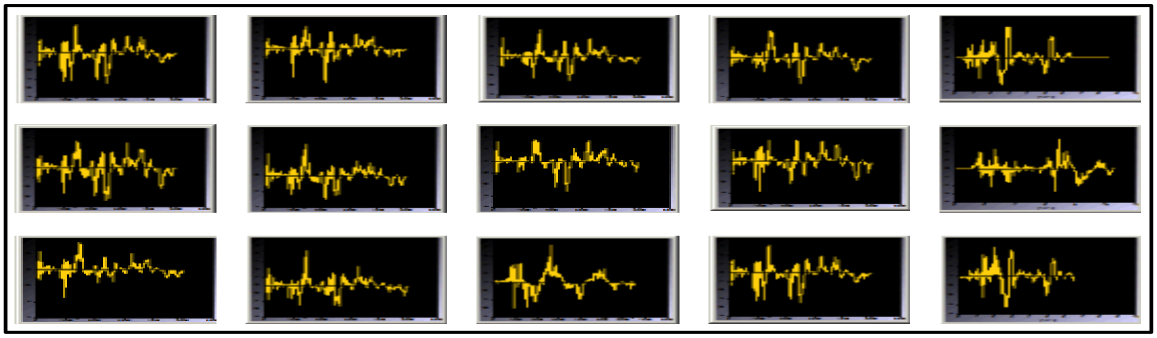


Figure 125: A Sample of Different Atomic Sound Events Generated from Closing of Sedan Door

The AEDCA system classifies different sound sources based on a tri-level scheme. At level one, the performance of the classifier is determined in terms of its capability for classification of the main sound source class or what we refer to as "detection focus" (i.e., sounds of objects, sounds of vehicles, sounds of humans). At this level, the objective is to determine how effectively the classifier will be able to discriminate one main class from other set of classes. For instance discriminate human generated sounds (e.g., walking or talking) from a set of other object sounds (e.g., vehicles or objects related sounds). At level two, the performance of the classifier is measured based on its performance for classification of the type of object, human, or vehicle associated with the sound. At this level, the objective is to determine how effectively the classifier will be able to determine the type of object associated with the sound source (e.g., metallic object, glass object, etc.). Finally, at level three, performance of the classifier is determined based on its capability to differentiate reliably between different types of interaction that generate the sound. For instance for a metallic object how will the classifier can determine if the sound source was generated due to a dropping of the metallic object on the ground or inside a vehicle trunk. Table 34, Table 35, and Table 36 illustrate performance comparison of both classifiers at different level of these tri-level stages. Based on the performance results from all levels, the CB-TM classifier was determined to yield superior performance over the k-d tree classifier by 2.8%.

Table 34: Level One Evaluation (Detection Focus)

	Object Sound	Vehicle Sound	Human Sound
CB-TM	76.8%	84%	86.2%
Kd-tree	73.3%	77.7%	87.1%

Table 35: Level Two Evaluation

	Object Material					Vehicle Type				Human State	
	Metallic	Wooden	Plastic	Glass	Ceramic	Sedan	SUV	Van	Truck	Walked	Talked
CB-TM	72.6%	84.2%	81.5%	78%	77.3%	80.9%	85.4%	80.9%	72.7%	86%	88%
Kd-tree	67.8%	81.5%	77.3%	76.2%	76%	79%	81.8%	77.2%	74.5%	84%	85%

Table 36: Level Three Evaluation

	HOI			HVI						Walking-Surface Type		
	Ground	Trunk	Carried	Door Opened	Door Closed	Hood Opened	Hood Closed	Trunk Opened	Trunk Closed	Concrete	Sand	Twigs
CB-TM	78.8%	75.2%	74.1%	80%	86%	86%	89%	89%	87%	87.5%	87.5%	75%
Kd-tree	73.5%	64.1%	74.1%	78%	86%	84%	88%	83%	87%	82.5%	86.2%	75%

3.4.4.9 Performance Evaluation of Activity Recognition

This section presents the performance evaluation of the HMM-AAR using different human activities such as phone/walkie-talkie conversation; loading, unloading, and moving of objects, and departure or arrival of vehicles. For computational aspect of the HMM-AAR, each event in the activity tested is numerically assigned a designated ID number as soon as its type and class is identified. Table 37 shows illustrative examples of three "Unloading Objects" activity cases where several atomic sequential events are recorded which signify the occurrence of unloading objects. The input to the HMM-ARR, are the numerically-index array of sequential events. For example, for the first unloading object case presented in the Table 37 this sequential indices include: 3→34→4.

Table 37: Sample of Different "Unloading Objects" Activity Cases

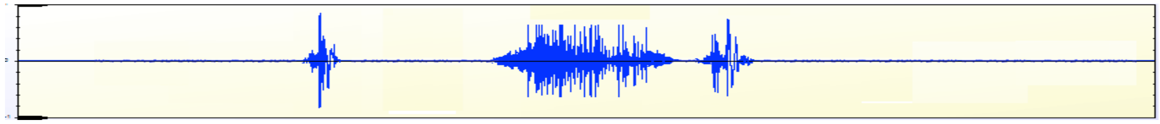
Unloading Objects -1 (Sequential Events):

Vehicle_Sedan_Trunk-Opened → Objec_Metallic_Dropped-Ground → Vehicle_Sedan_Trunk-Closed

Unloading Objects -1 (Corresponding Coding IDs):

3 → 34 → 4

Unloading Objects -1(Corresponding Sequential Sounds Events):



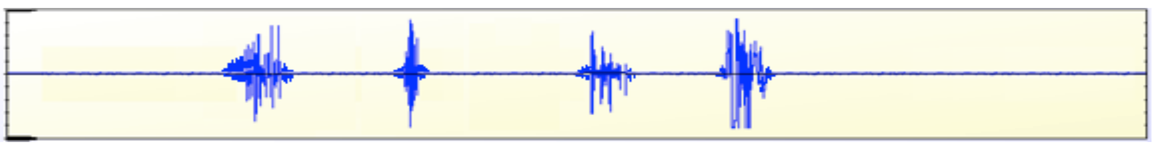
Unloading Objects -2(Sequential Events):

Vehicle_Sedan_Trunk-Opened → Object_Ceramic_Dropped-Ground → Object_Plastic_Dropped-Ground → Vehicle_Sedan_Trunk-Closed

Unloading Objects -2 (Corresponding Coding IDs):

3 → 38 → 28 → 4

Unloading Objects -2(Corresponding Sequential Sounds Events):



Unloading Objects -3(Sequential Events):

Vehicle_SUV_Trunk-Opened →Object_Plastic_Dropped-Ground → Object_Wooden_Dropped-Ground → Objec_Metallic_Placed-Ground → Vehicle_SUV_Trunk-Closed

Unloading Objects -3 (Corresponding Coding IDs):

10 → 28 → 31 → 34 → 4

Unloading Objects -3(Corresponding Sequential Sounds Events):

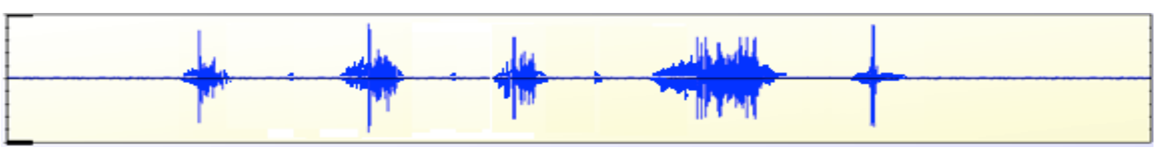


Table 38 demonstrates the results of the HMM-AAR using fifteen experimented cases with different sounds context. The numbers under Count Column, in the Table 38, corresponds to the number of test ground-truth cases considered. The number under the "Correctly Recognized Cases" is the number of the correctly classified cases.

Table 38: HMM-AAR Performance Measure

	Count (Ground-truth cases)	Correctly Predicted Cases	Incorrectly Predicted Cases	Unknown Case
Phone Conversation Activity	15	12	1	2
Walkie-Talkie Conversation Activity	15	13	0	2
Loading Objects Activity	15	11	1	3
Unloading Objects Activity	15	10	2	3
Moving Objects Activity	15	11	0	4
Vehicle Arrived Activity	15	12	0	3
Vehicle Departed Activity	15	13	0	2
Total	**105**	**82**	**4**	**19**

Based on the total result presented in the Table 38, from the 105 different trained cases tested, the HMM-AAR was able to predict 82 cases correctly, 4 cases were recognized incorrectly, and 19 cases were misrecognized. Based on the performance results the Activity Recondition shows a 78% overall accuracy. In the following section, a method developed in AEDCA for generating semantic annotations from recognized acoustic events and activities is described in details.

3.4.4.10 Acoustic Sensors Semantic Annotation

A sound event or activity can be annotated by certain key linguistic term. Some events are, in general, context-sensitive. Therefore, a care should be taken into account to ensure the correct terminology applied for expressing recognized events. The role of semantic message generation is to convert an acoustic observation to a structured linguistic grammar understandable to people. Another aspect of an automatic audio scene recognition system is to provide automatic annotation of sound events without direct involvement of human supervisor in the loop. One way to semantically annotate acoustic data is through a structured Transducer Markup Language (TML) [J.22] which is a schema for capturing, characterizing, and enabling sensor meta-data reporting. To effectively describe the acoustic perception of human interaction in the environment, it suffices to semantically describe such interactions based on two known states: (1) event state and (2) activity state. Figure 126 illustrates how the AEDCA manages to extracts attributes from different sound events and associated each detected attributed with a certainty measure. The certainty measures are obtained from historical performance evaluation of the AEDCA atomic sound classifier.

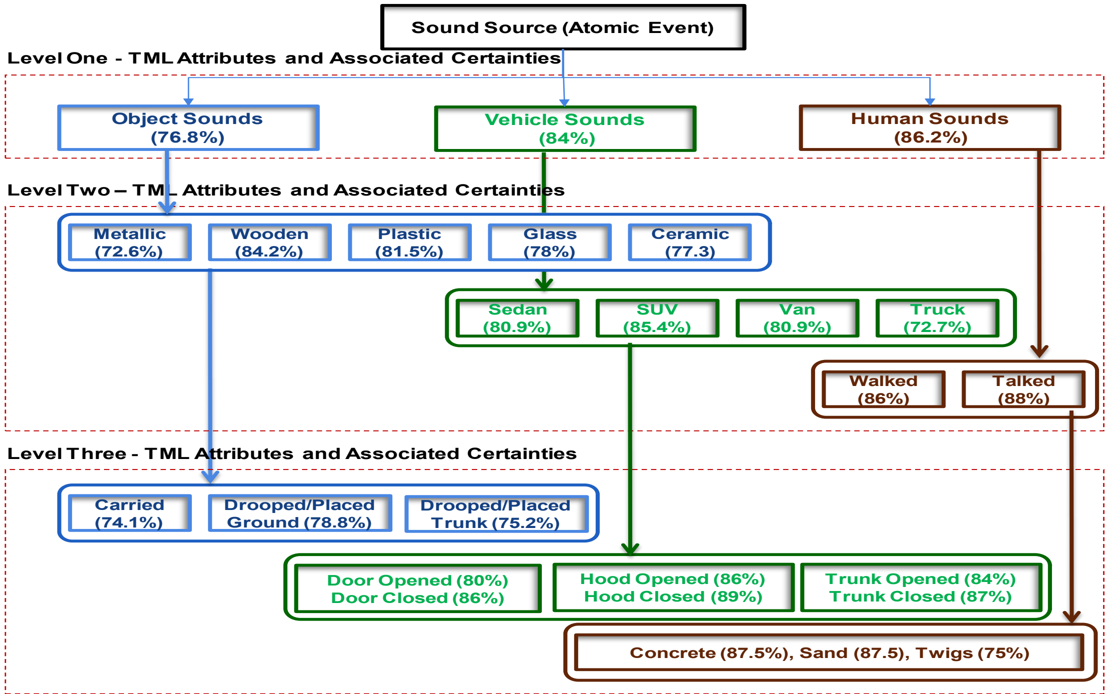


Figure 126: Certainty Associated with Detection of Sound Attributes for different categories of interactions (i.e., HHI, HVI, HOI)

3.4.4.11 Event State Annotation

In the event state, the most aggregated information is employed for annotating the recognized atomic acoustic events. Listed below is an example of the TML meta-data format developed for describing acoustic events:

<data_ref="source">yyyy-mm ddT,hh:mm.ss,Message_ID,Global_Space,Source_Attribute,Location_Space, Detection_Focus,Detection_ID,Attribute-1;Attribute-2;Attribute-3;Attribute-4;Attribute-5;Attribute-6;Group-ID;Confidence</datd_ref="source">

As demonstrated, this TML format offers a multi-facet scheme for tagging different attributes of a recognized event. This format begins with the *date* and *time* tags representing the actual date and time an event is detected respectively. The date is expressed in format of yyyy-mm-dd, namely, designating the year, month, and day that the event was detected. Similarly, the format hh:mm.ss designate the time format for express the actual hour, minute, and second where the event was reported by the sensor. The following *Message-ID* tag denotes the sequential number of the generated message. The *Global_Space* tag denotes the location of the source specified in form of GPS coordinates. The *Source_Attribute* tag denotes a specific sensor setting if applicable. The *Location_Space* tag denotes a logical name used to describe the monitored environment location (e.g., parking lot, warehouse). The *Detection_Focus* tag denotes the event of interest (e.g., human, object, or vehicle), and the ordered list of Attributes 1 through 4 denotes the type, content, state, and interaction of the detection focus respectively. The Confidence tag

demonstrates the confidence level calculated using the certainty associated with each attribute in the TML message. Let Conf = $\{\mathrm{Conf}_1, \mathrm{Conf}_2, \ldots\ldots \mathrm{Conf}_n\}$,then:

$$Avg_{conf} = (Conf_1 + \cdots Conf_n)/n$$

Where n represents total number of sound attributes reported. The example given below illustrates a typical generated TML after recognizing a human interaction with a metallic object in the environment:

<Acoustic=Acs-1>2012-08-16T,13:22.05,05,33.3009,44.3949,Warehouse_1,Zone-1,Object,None,None, Metallic,None,None,Carried ,None,None,74.5</Acoustic=Acs-1>

The final confidence level of the generating TML message is specified using an average of confidence levels of all related attributes as presented in the Figure 126. The confidence level in this TML message is determined as an average of (1) the certainty related to detection of the sound source as an object (i.e., 76.8%), (2) the certainty related to determining the type of the object as a Metallic (i.e., 72.6%), and (3) the certainty associated the operation generated the sound which was carrying the object in this case (i.e., 74.1%)).

3.4.4.12 Sound Activity Annotation

In the activity state, on the other hand, the semantic messages are generated when the confidence threshold of an identified activity exceeds a threshold limit. Listed below is an example of a TML format generated for describing the human sound activity that are taking place in the environment:

<data_ref="Acoustic">yyyy-mm-dd,Shh:mm.ss,Ehh:mm.ss,Message-ID, Global_Space, Location_Space, Activity_Name, Messages_ID;Confidence </datd_ref ="Acoustic">

As demonstrated, this format begins with the date tag representing the actual date an activity is detected. The starting time format Shh:mm.ss designate the actual hour, minute, and second where the activity was started. Similarly, the ending time format Ehh:mm:ss designate the actual hour, minute, and second where the activity was ended. The following Message-ID tag denotes the sequential number of the generated message. The Global_Space tag denotes the location of the source specified in form of GPS coordinates. The Location_Space tag denotes a logical name used to describe the monitored environment location (e,g., parking lot, warehouse). The Activity_Name denotes the name of the activity recognized (e.g., loading activity, unloading activity). The Messages_ID tag denotes a reference to all messages associated with the detected activity, and finally, the Confidence tag is an aggregated confidence level of all events messages involved in this activity. The example given below illustrates a typical generated TML message to effectively describe an "Unloading Objects" activity:

<data_ref="Acoustic">2014-03-30,S13:22.24,E13:23.46,96,33.30094,44.39491,Market_place ,Zone-1, Unloading Objects,70-71-72-73-74-75;83.5 </data_ref =" Acoustic">

3.4.4.13 Final Remarks on Acoustic Signal Processing and Annotation

For the scope of this project, we developed an approach for semantic labeling of acoustic signatures describing different human interactions and activities happening in the environment. In our approach, sound events are initially segmented into a set of atomic events. Two classifiers, namely, CB-TM and k-d trees were employed for classification of atomic acoustic events. In the event classification stage three level of evaluation were considered (i.e., detection focus, type, and interaction). Appropriate numbers of individual sound events were used at each level for training and testing the performance of the proposed approach. At the event classification stage the CB-TM was superior to the k-d trees in all performance metrics. We believe that the reason behind this result is because the CB-TM is less sensitive to variations than the k-d tree. However, the renowned advantage of k-d trees over the CB-TM that they are faster to query. They also require less memory, since they consist of a single tree rather than an ensemble. At the Activity Recognition stage the developed HMM-AAR was evaluated using different activities cases such as phone/walkie-talkie conversation; loading, unloading, and moving of an objects; and departure or arrival of vehicles. It was observed through this examination that the performance of the HMM-AAR is coupled with the performance of the atomic event classifier. Finally, the role of semantic message generation (i.e., event state and activity state) was demonstrated to convert an acoustic observation to a structured linguistic grammar understandable to people. The results strongly suggest that the approach is both reliable and robust, and can be extended to future PSS applications.

3.4.5 Decision Support System

As demonstrated earlier, prior to fusion of hard and soft data, we resort to semantically annotate hard sensor data. Human observational reports, generally, represent body of texts describing situation of entities, events, and actions taking place under different operational circumstances. In general, integration of hard and soft sensor sources is meaningful if they are spatiotemporally correlated and there exists a strong association between the two sources and they both support a shared concept. When large amount of soft and hard sensory sources are available, there would be a need for an intelligent decision support system that can perform automatic text mining and reveal all soft and hard sensor sources supporting a given concept query.

The decision support system we have developed under this study is based on the proven Latent Semantic Analysis (LSA) technique - a fully automatic mathematical/statistical technique for extracting and inferring relations of expected contextual usage of words in passages of discourse. LSA is not a traditional natural language processing or artificial intelligence program; it uses no humanly constructed dictionaries, knowledge bases, semantic networks, grammars, syntactic parsers, or morphologies, or the like, and takes as its input only raw text parsed into words defined as unique character strings and separated into meaningful passages or samples such as sentences or paragraphs. To apply LSA, we primarily treat each data associated with the hard sensor (i.e., TML annotated messages) and each observer's report as a separate document so that they can be indexed.

Next, our LSA applies singular value decomposition (SVD) to the matrix. This is a form of factor analysis, or more properly the mathematical generalization of which factor analysis is a special case. In SVD, a rectangular matrix is decomposed into the product of three other

matrices. One component matrix describes the original row entities as vectors of derived orthogonal factor values, another describes the original column entities in the same way, and the third is a diagonal matrix containing scaling values such that when the three components are matrix-multiplied, the original matrix is reconstructed.

As a practical method for the characterization of word meaning, we know that LSA produces measures of word-word, word-passage and passage-passage relations that are well correlated with several human cognitive phenomena involving association or semantic similarity. Empirical evidence of this will be reviewed shortly. The correlations demonstrate close resemblance between what LSA extracts and the way peoples' representations of meaning reflect what they have read and heard, as well as the way human representation of meaning is reflected in the word choice of writers. As one practical consequence of this correspondence, LSA allows us to closely approximate human judgments of meaning similarity between words and to objectively predict the consequences of overall word-based similarity between passages, estimates of which often figure prominently in research on discourse processing.

In general, information fusion systems are intended, most of the time, to be used by human analysts to assist them with their decision making. Therefore, the objective of a fusion system should be tailored towards supporting decision-makers. This process is involved with a computer interface that ultimately presents to the analysts the information resulting from the data fusion process. Consequently, an information fusion system could, indeed, be treated as DSS. When considering Information Fusion System (IFS) as DSS there are a number of benefits. First of all, information fusion as currently constituted is lacking *a user perspective*, treating the IFS as a DSS would naturally give this perspective. Secondly, enabling this user perspective and treating information fusion system as DSS may *ensure the effectiveness of the system*. Today, information fusion system become more and more advanced and data-intensive and in order to optimize the system it is no longer acceptable to just use more advanced sensors for lack of having cognitive understanding of the data. The ultimate performance of a decision support system does not only depend on its quality, but also of the possible utilization of the system in practice. Therefore, it is clear that there is a compelling reason to use information fusion from another perspective, i.e. the *Decision support perspective*. Thirdly, treating information fusion system as a DSS could give a top down perspective, demanding more high level research toward decision fusion. In other words, having a top-down perspective would give the focus of who should be supported by the system, what information we need in order to make the decision. This would enable the needs and requirements of users to be considered before considering what data we could actually fuse and what relationships we could find in the data to support a decision. Figure 127 illustrates results of a concept query from a collection of soft and hard messages/documents. The query is composed of three words, "***loud noise at night***". In PLSA, the order of words describing the Concept is immaterial, so are the stop words (e.g., at as appears in the query statement). As shown, the decision support system automatically ranks 10 top most relevant messages/documents with a probability measure reflecting their degree of semantic closeness to the specified Query. As illustrates, no single documents in the database completely support this query.

3.4.5.1 Probabilistic Latent Semantic Analysis

Probabilistic Latent Semantic Analysis (PLSA), also called Probabilistic Latent Semantic Indexing (PLSI), is based on statistical Latent Semantic Analysis (LSA) technique that is established based on a theory and method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text. PLSA or PLSI is a novel approach to automated document indexing which is based on a statistical latent class model for factor analysis of count data. Fitted from a training corpus of text documents by a generalization of the Expectation Maximization algorithm, the utilized model is able to deal with domain specific synonymy as well as with polysemous words.

The underlying technique behind PLSA is the Latent Semantic Analysis (LSA). LSA is an approach to automatic indexing and information retrieval that attempts to overcome these problems by mapping documents as well as terms to a representation in the so called latent semantic space. LSA usually takes the (high dimensional) vector space representation of documents based on term frequencies as a starting point and applies a dimension reducing linear projection. The specific form of this mapping is determined by a given document collection and is based on a well-known technique called Singular Value Decomposition (SVD) that forms a corresponding term/document matrix. The general intuitiveness of this approach is that similarities between documents or between documents and queries can be more reliably estimated in the reduced latent space representation than in the original representation.

LSA power is realized in terms of its ability to serve as a natural language processing tool, in particular in vectorial semantics, and for analysis of relationships between a set of documents and the terms they contain by producing a set of concepts related to the documents and terms. LSA assumes that words that are close in meaning will occur in similar pieces of text. A matrix containing word counts per paragraph (rows represent unique words and columns represent each paragraph) is constructed from a large piece of text and a mathematical SVD technique to reduce the number of columns while preserving the similarity structure among rows. Words are then compared by taking the cosine of the angle between the two vectors formed by any two rows. Values close to 1 represent very similar words while values close to 0 represent very dissimilar words.

Let the document collection be represented by a (D,n) matrix $X = |x_1|, \ldots, |x_n|$, where the columns are document Bag-of-Words vectors, D is the vocabulary size, and n is the number of documents. LSI is SVD applied to X:

$$X_{Dxn} = U_{D\times m} S_{m\times m} V_{n\times n}^T \tag{1}$$

Furthermore, one only keeps the largest $d \leq m = min\ (D,\ n)$ singular values. That is, let $\widehat{U}_{D\times d}$ be the first d columns of U, $\hat{S}_{d\times d}$be the first $d\times\ d$ submatrix of S, $\hat{V}_{n\times d}$ be the first d columns of V . Then

$$\widehat{U}_{D\times d} \hat{S}_{d\times d} \hat{V}_{n\times d}^T \tag{2}$$

is the best rank-d approximation to X in the least square sense. The d columns of $\widehat{U}_{D\times d}$ defines the new rotated lower dimensional coordinate system. The n columns of $\hat{S}_{d\times d} \hat{V}_{n\times d}^T$ are the new coordinates of each document after dimensionality reduction. The new coordinate system allows a natural way to perform dimensionality reduction for points not in the dataset. For a new test document x^*, its new coordinate is $\widehat{U}_{x^*}^T$. If x^* coincides with an existing document

x_i, it will have the same new coordinates $\widehat{U}_{x_i}^T = \hat{S}\hat{V}_i^T$ as above. An alternative would be to add x^* to the dataset, and compute SVD on the $n+1$ documents. In general, this will produce a slightly different dimension reduction solution. It is much more computational intensive.

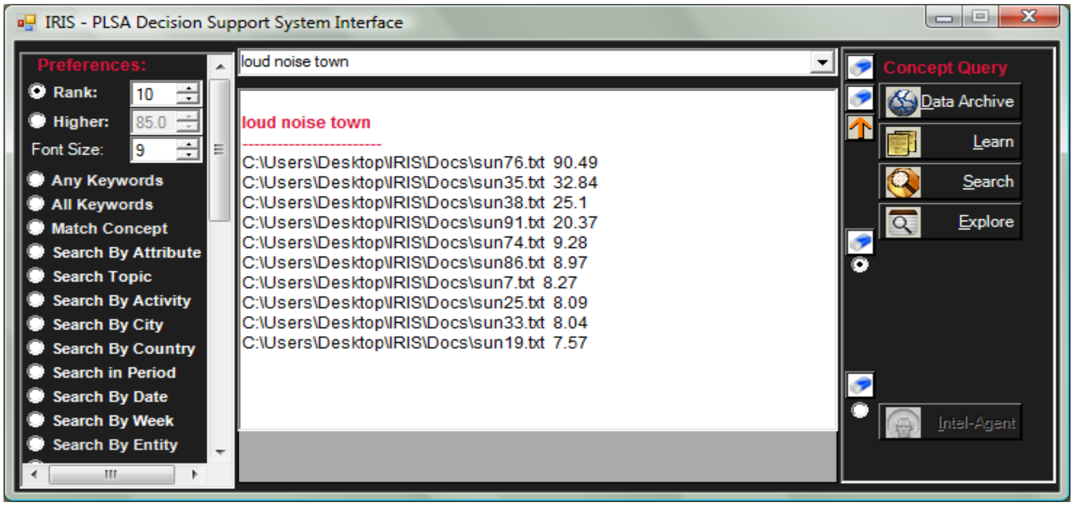


Figure 127: Intelligent Decision Support System Interface

The core of PLSA is a statistical model which has been called aspect model [J.12]. The latter is a latent variable model for general co-occurrence data which associates an unobserved class variable $z \in Z = \{z_1, z_2 \ldots\ldots, z_k\}$ with each observation, i.e., with each occurrence of word $w \in W = \{w_1, w_2 \ldots\ldots, w_m\}$ in a document $d \in D = \{d_1, d_2 \ldots\ldots, d_N\}$. In terms of a generative model it can be de_ned in the following way:

- Select a document d with probability $P\,(d)$,
- Select a latent class z with probability $P\,(z|d)$,
- Generate a word w with probability $P\,(w|z)$.

As a result one obtains an observed pair (d;w), while the latent class variable z is discarded. Translating this process into a joint probability model results in the expression:

$$P\,(d, w) = P\,(d)\,P\,(w|d) \tag{3}$$

where

$$P\,(w|d) = \sum_{z \in Z} P\,(w|z) P\,(z|d) \tag{4}$$

Following the likelihood principle, one can determine $P(d)$, $P\,(z|d)$, and $P\,(w|z)$ by maximization of the log likelihood function:

$$l = \sum_{z \in Z} \sum_{z \in Z} \tau(d, w) \log P\,(w, d) \tag{5}$$

where $\tau(d, w)$ denotes the term frequency, i.e., the number of times w occurred in d. It is worth noticing that an equivalent symmetric version of the model can be obtained by inverting the conditional probability $P\ (z|d)$ with the help of Bayes' rule, which results in:

$$P\ (d, w) = \sum_{z \in Z} P\ (z) P\ (w|z) P(d|z) \tag{6}$$

The rationale is that documents which share frequently co-occurring terms will have a similar representation in the latent space, even if they have no terms in common. LSA thus has potential to perform some sort of noise reduction and has the potential benefit to detect synonyms as well as words that refer to the same topic. In many applications this has proven to result in more robust word processing [J.9][J.10][J.11]. In this paper, we applied Probabilistic Latent Semantic Analysis (PLSA) that offers a solid statistical foundation since it is based on the likelihood principle and defines a proper generative model of the data. In brevity of space limitation, we avoid describing PLSA technique in details here and encourage interesting readers to refer to [J.9] for an excellent description of this technique.

Secondly, we represent the text as a matrix in which each row stands for a unique word and each column stands for a text passage or other context. Each cell contains the frequency with which the word of its row appears in the passage denoted by its column. Next, the cell entries are subjected to a preliminary transformation, in which each cell frequency is weighted by a function that expresses both the word's importance in the particular passage and the degree to which the word type carries information in the domain of discourse in general.

Figure 128 illustrates the content of the top three messages/documents that support this requested concept query. The keywords in each document that support the query partially are automatically highlighted by the decision support system. The DSS model is able to search the corps of sensor sources (i.e., messages/documents) automatically when in PLSA mode. It can also search sensor sources based on a number of user specified key attributes. These attributes include: Search by Attribute, Topic, Activity, City, Country, Period, Date, Week, Entity, GPS Coordinates, Name, Job, Friend, Relative, Associate, Place, Streets, Object, Phone #, Affiliation, Tools, Clothing, Event, Beliefs, Network, and Bio Information. The former search technique search corps of sensor sources semantically, however, the latter search techniques, individually, or in a collection can further search corps of existing documents to the specific information the analyst user had requested.

iDSS offers an Intelligent Agent that recommends the next most relevant key to the Concept Query the maximize the entropy of semantic relationship among most selected sensor source messages/documents. As shown in Figure 129, the Intelligent Agent recommends a new keyword "D_A", an abbreviated name. Based on this new Concept Query, a new set of search results are obtained as illustrated in Figure 129 and Figure 130.

One of the key challenges that we have encountered is the development of data formats, protocols, and methodologies to establish an information architecture and framework for the effective capture, representation, transmission, and storage of the vastly heterogeneous data and accompanying metadata -- including capabilities and characteristics of human observers, uncertainty of human observations, "soft" contextual data, and information pedigree. In this paper, we presented an intelligent decision support system (*iDSS*) with capability to either semantically and/or explicitly search large corps of soft/hard messages stored in the database in

form of separate documents. The *iDSS* takes advantage of the power of probabilistic latent semantic analysis (PLSA) technique to achieve its former capability, and apply an information technology (IT) technique to search soft/hard messages/documents supporting specific explorations.

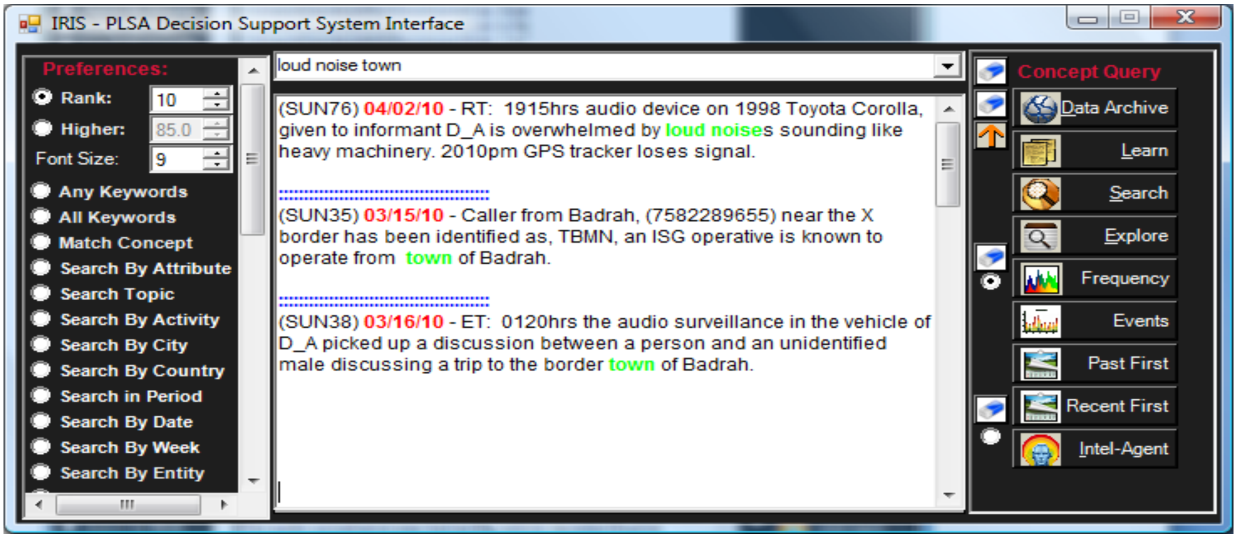


Figure 128: Results of the Intelligent Decision Support System

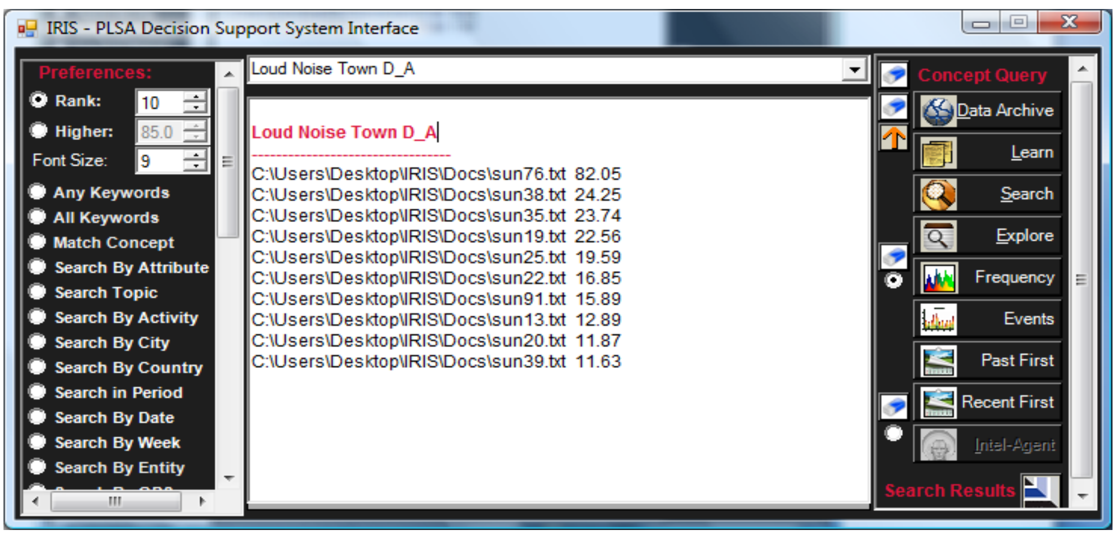


Figure 129: Results of the Intelligent Decision Support System (After Second Concept Query Iteration)

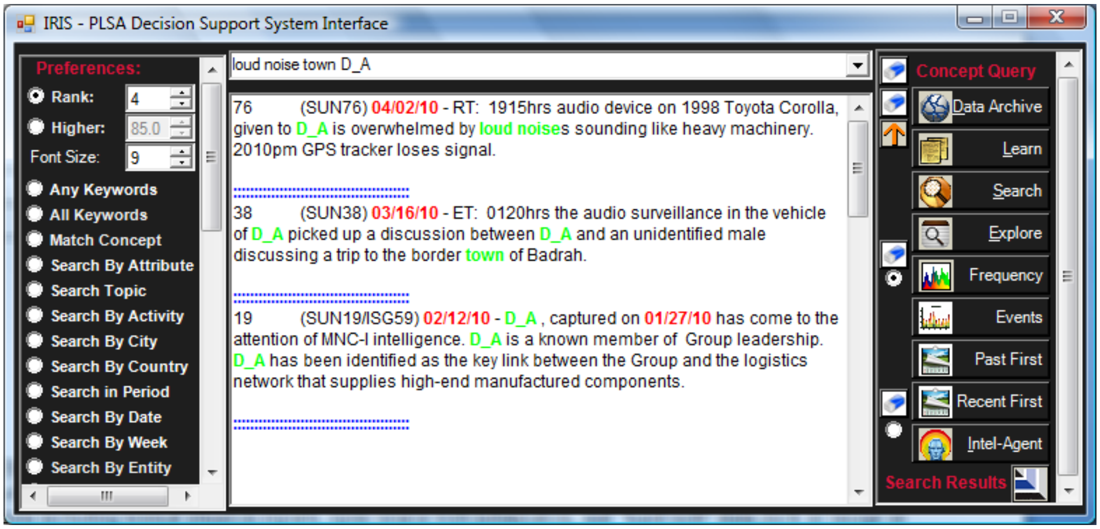


Figure 130: Results of the Intelligent Decision Support System (After Second Concept Query Iteration)

3.4.6 Experimental Research Work

In this period, we also conducted a series of emulated SYNCOIN experiments with human-in-the-loop. The context of scenarios was selected based on the synthetic SUNNY message set created by the Pennsylvania State University (PSU). Each experiment was conducted by participation of two or more operators. The scenarios were conducted in urban environment and video recorded via two or more remotely located PTZ cameras. For each experiment, both imagery and acoustic data were collected. The sensor data from each experiment were processed and corresponding TML messages for each experiment was generated for integration with other hard/soft sensor fusion activities of this MURI project. A collection of outdoor experiments conducted at TSU in this period are presented in Figures Figure 131 through Figure 144. In addition to optical camera, we used Kinect depth map cameras for characterization of human-vehicle interaction in outdoor environments. The results of these experiments are detailed in our references [J.1]-[J.5].

Figure 131: A HVI unloading group activity participated by four-operators

Figure 132: A social network activity involved with vehicles exchange

Figure 133: A group activity emulating parts loading from a warehouse

Figure 134: A social network involved with three persons

Figure 135: Another social network activity with some background noise

Figure 136: A four-operator group activity demonstrating a social network negotiation

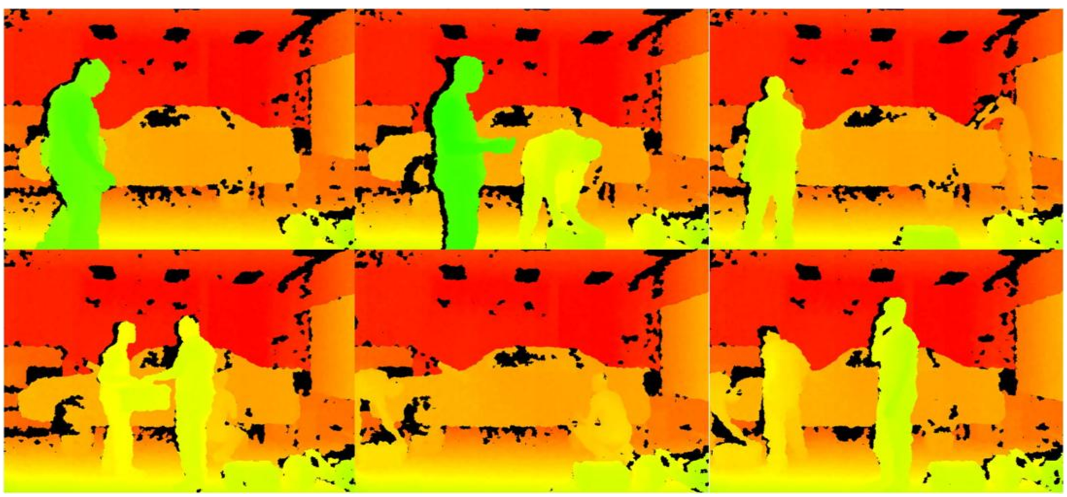


Figure 137: A group activity involved with parts loading of boxes of different sizes

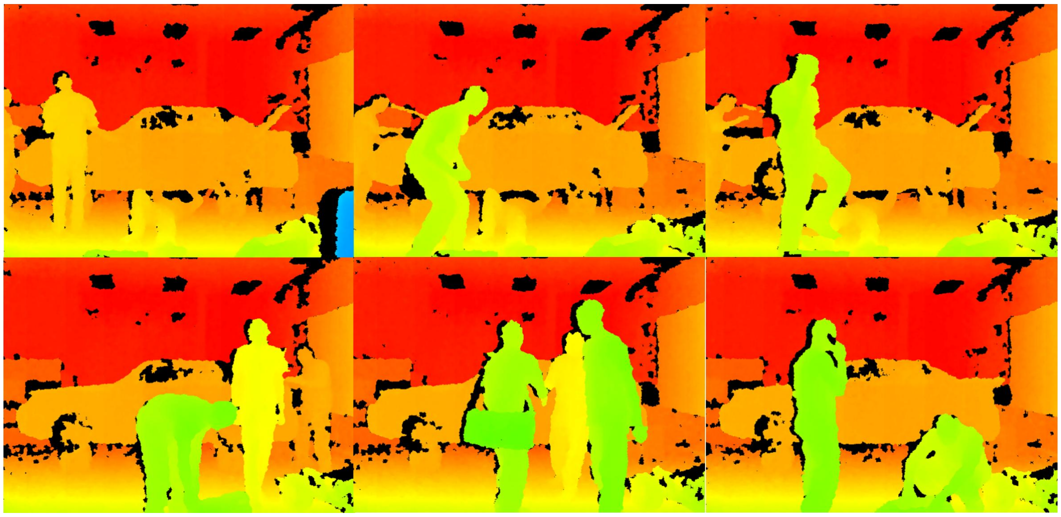


Figure 138: A group activity involved with parts delivery

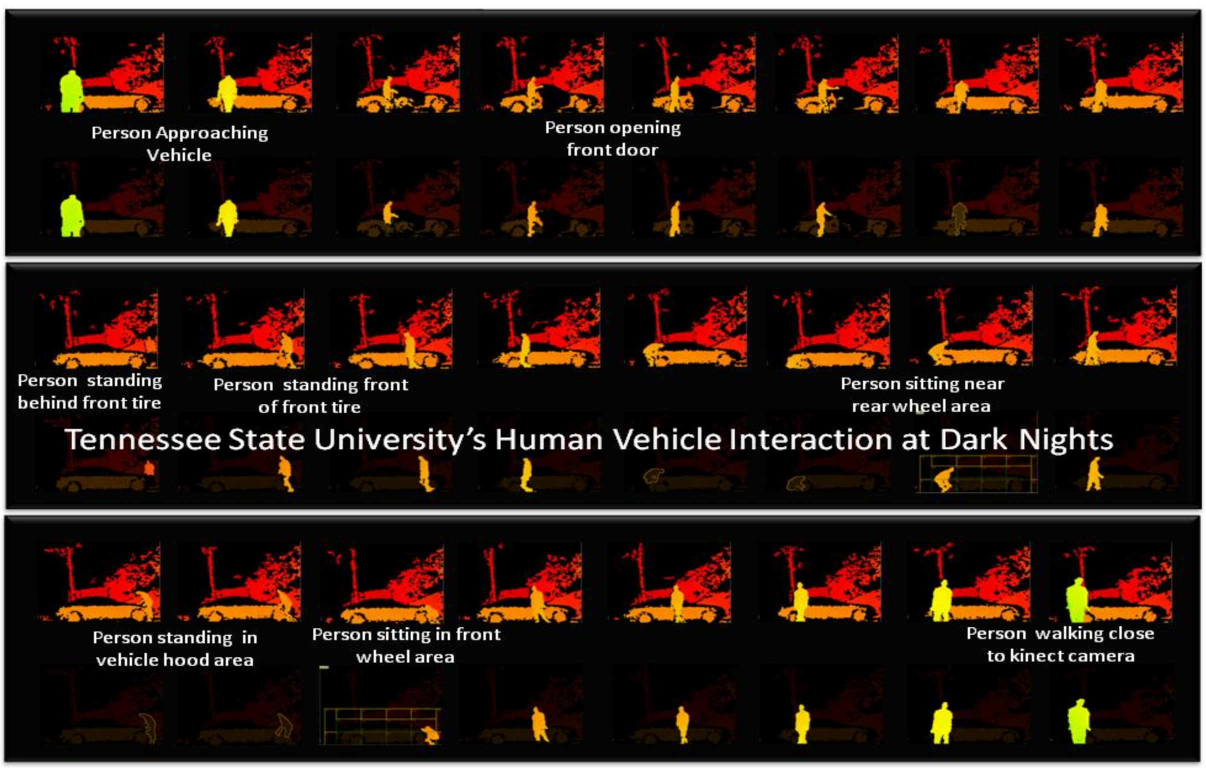


Figure 139: An outdoor experiment conducted for Human-Vehicle Interaction (HVI) detection with Kinetic Range Map Camera

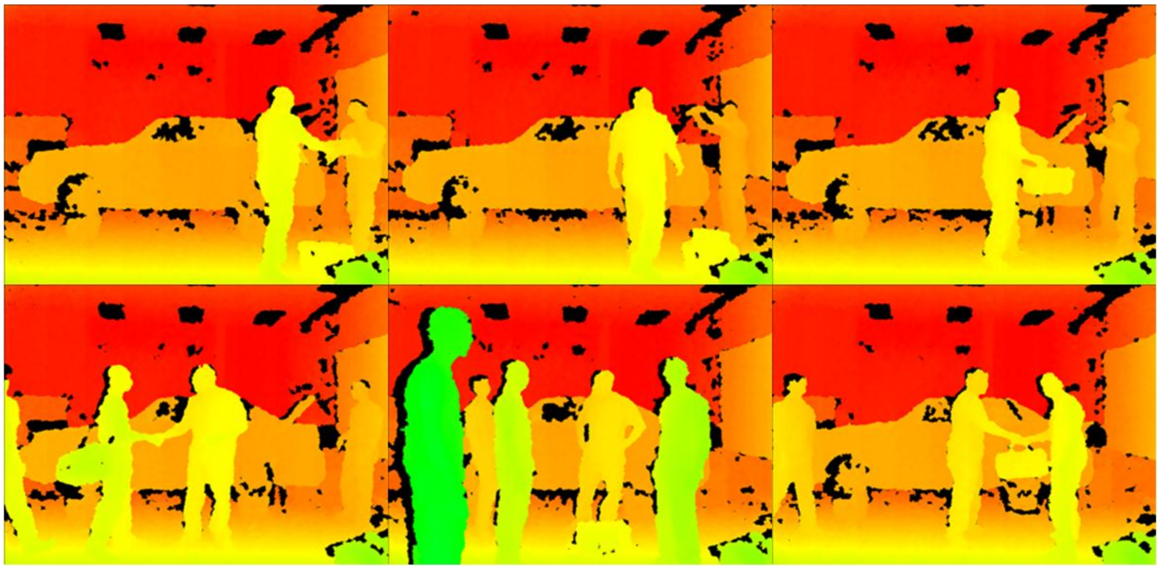


Figure 140: A social network activity involved with parts delivering, trunk loading, and negotiation

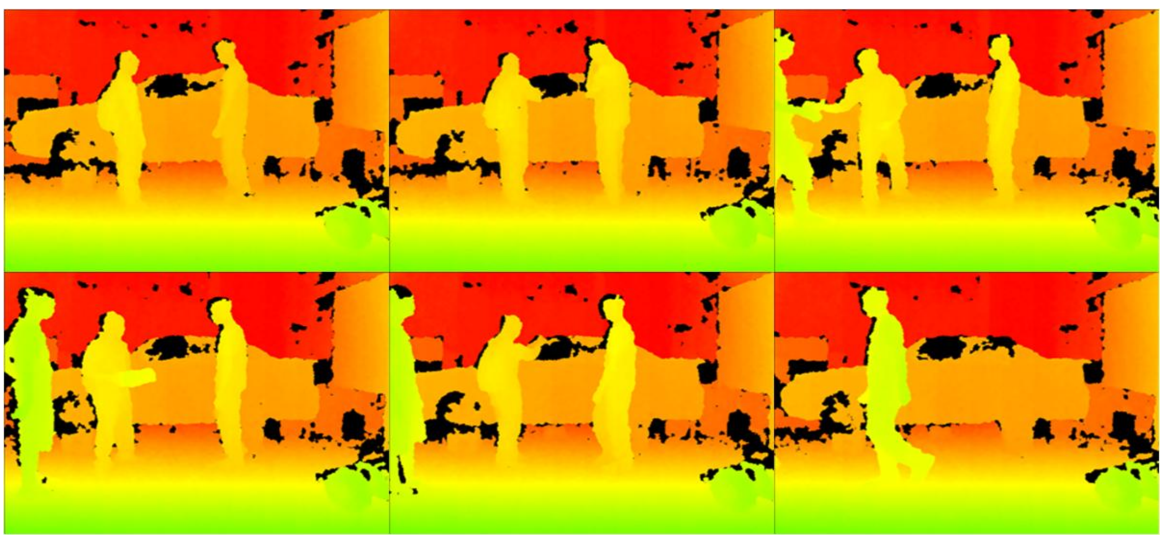


Figure 141: A social network involved with the driver inside of the vehicle playing the role of a leader

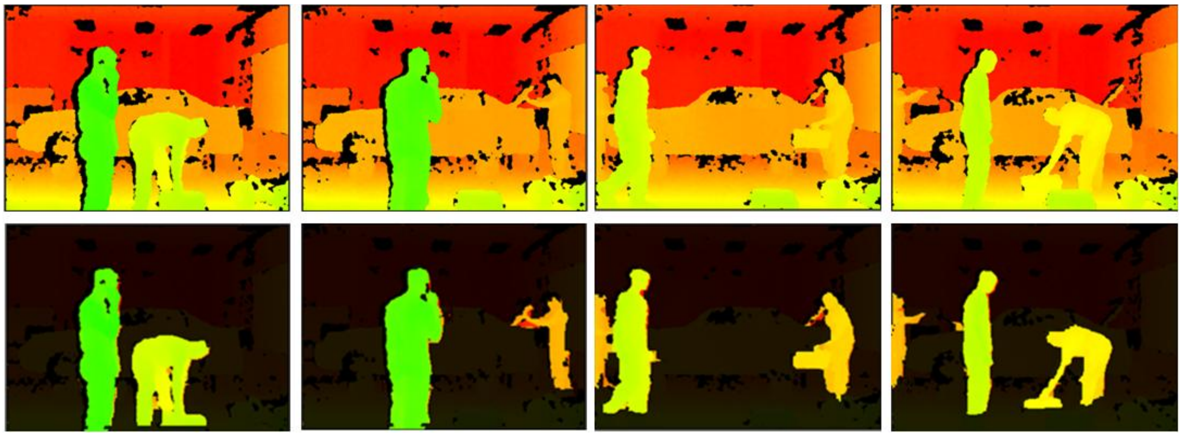


Figure 142: Image processing results of Kinect depth map camera

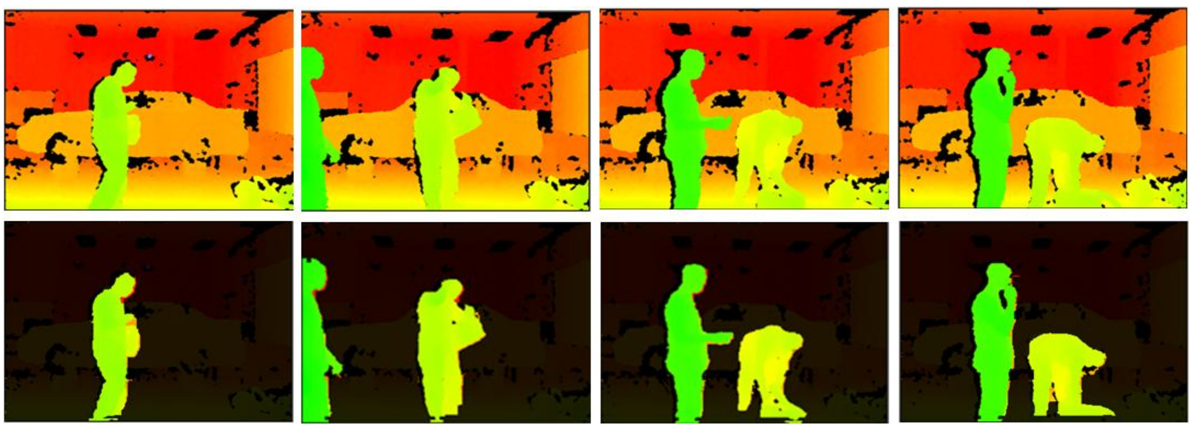


Figure 143: Image processing results of Kinect depth map camera

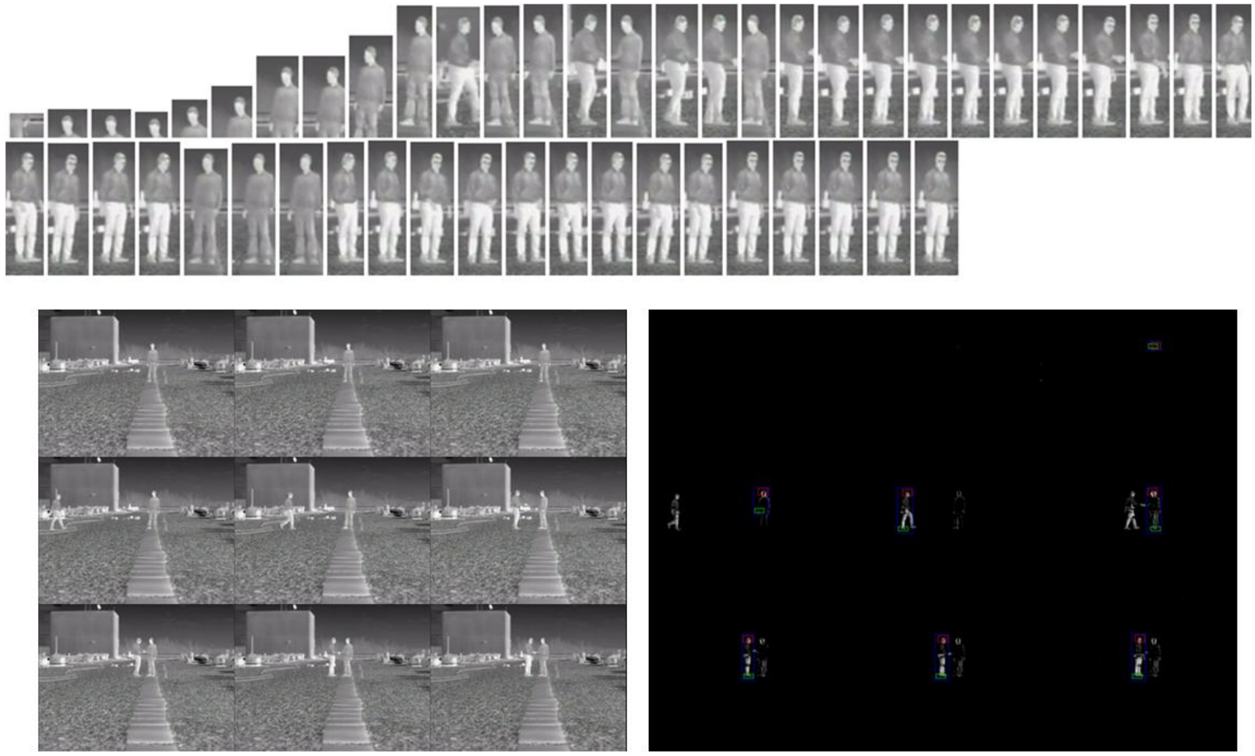


Figure 144: Image processing results from IR camera

3.4.7 Knowledge Discovery in Group Activities: Sequential Observation Analysis

Understanding of group activities is complex and essential for intelligent PSS applications. To enable understanding of such a complex activities it requires identification of actions and interactions at different level of data abstraction. An archetypal group activity may compose of one or more distinctive interactions that we define them as: [J.12]:

Human-Human Interactions (HHI): this category of interactions deals with identification of casual interactions individuals in the group may have with others and with other members of other groups they are associated with. Understanding of such interactions may reveal role of individuals in the group and may explain their nature of interactions. Examples of HHI events are: shaking hands, greeting, talking, cooperating, co-walking, etc.

Human-Vehicle Interactions (HVI): this category of interactions deals with types of interactions that one may have with a vehicle used in an activity. Such interactions may include opening-closing doors/trunk/hood, driving/parking, and etc. Vehicles have been used as a primary source of transportation for pursuing many outdoor suspicious activities [J.12]. Analysis of the Human-Vehicle Interactions (HVI) can lead to identify cohesive patterns of activities representing potential threats. Identification of such patterns can significantly improve situational awareness in intelligent PSS.

Human-Object Interactions (HOI): this category of interactions consist of type of interaction(s) that one may have with respect certain object(s) used in a group activity. These types of interactions are typically difficult to detect by the nature of object(s) involved. This category of interactions represents the most challenging category for analysis. Examples of HOI

events are: person carrying an object, person dropping an object, person placing an object in vehicle, etc. Due to inherence obstruction and occlusion involved with detection of such events, it becomes a difficult challenge to detect and verify such events reliably.

Behavioral characterization of targets of interest that follow a prior known ontology may lead to the discovery of nature of a GA. For the objective of this paper, we differentiate among three terms: "Action", "Activity", and "Group Activity".

An *Action* refers to a simple operational patterns that a single or multiple entity(ies) may perform to alter that state of the environment. (e.g., a person opening a vehicle's trunk, group walking, person dropping object etc.).

An *Activity* refers to an array of associated and correlative actions that an entity may collectively perform to fulfill a specific task objective (e.g., a person loading object to a vehicle's trunk, a person changing vehicle's tire etc.). In general, we can express as [*Action(1) - Action(2) - Action(3) -- Action(n)* => Activity(x_1)] Example: 'Loading object in a vehicle' actions sequence: [*Person_carrying_a_object Person_opens_vehicle_trunk Person_placing_object_in_trunk* => Loading Object in vehicle]

A *Group Activity* refers to an array of spatiotemporally coupled activities that a group of entities may perform to achieve a specific common task objective (e.g., a team of individual cooperating for unloading a heavy object from a vehicle).

By proper characterization of such interactions, appropriate semantic messages can be generated to describe the attributes of activities taking place with their spatiotemporal significance. Such characterizations are considered vital to fusion of multi-modality sensor data. Semantic annotation of sensor observations can be performed in abstract sense after all and part of interactions are realized or be performed per each sensor observation that reveals a new piece of information. With the latter situation, the information embodiment of each semantic message is partial [J.12]. In the former situation, the challenge is that of comprehending each observation, associating information pertaining to each frame properly, and applying a consistent format for annotating attributes of sequential observations. In this paper, we endeavor the latter approach since it facilitates inference of group activities with a better traceability than the former approach as well as accommodating annotation of multi-modality sensor using a consistent data structure format. We model the GA recognition problem using HMM through sequential imagery observations and compare performance of three competing HMM architectures for recognition of group activities under different operational conditions. The remaining part of this paper is organized as such: group activity discovery and recognition framework, HMM modeling architectures, results analysis, and conclusion followed by acknowledgements and references.

3.4.7.1 Group Activity Discovery and Recognition (GADR) Framework

Group activity detection in PSS starts with the detection, identification, and tracking of targets via processing of data from multi-modality sensors (e.g., surveillance cameras, acoustic sensors or any other active monitoring sensors). On this project, we mainly focused on imaging cameras as the prime source for data. In brevity of space limitation here, hereon we make an assumption that sequential images from each video stream are processed and target of interests are detected and identified. The scope of this paper is limited to the recognition of GA using different architectures of HMM. More details about the Targets detection, identification, GA

discovery and Visual analytics can be found from our previous publications [J.1]-[J.3], [J.15]-[J.18]. Figure 145 shows an experimental example of a GA taking place in a parking lot near a safe house involving combinations of HVI, HHI and HOI. The tested GA scenario is named as "Loading objects at safe house" contained the following sequence of interactions "Vehicle arrives at a safe house"➔"A social network takes place"➔"Objects are removed from the Vehicle"➔"Objects are loaded into the safe house"➔ "Vehicle leaves the safe house". Figure 146 presents a six-stage of a Group Activity Monitoring System.

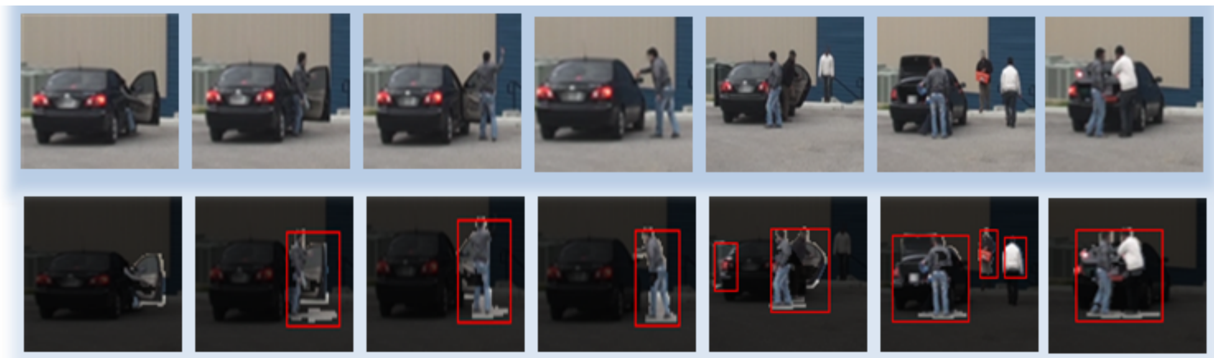


Figure 145: "Loading objects at safe house" Scenario

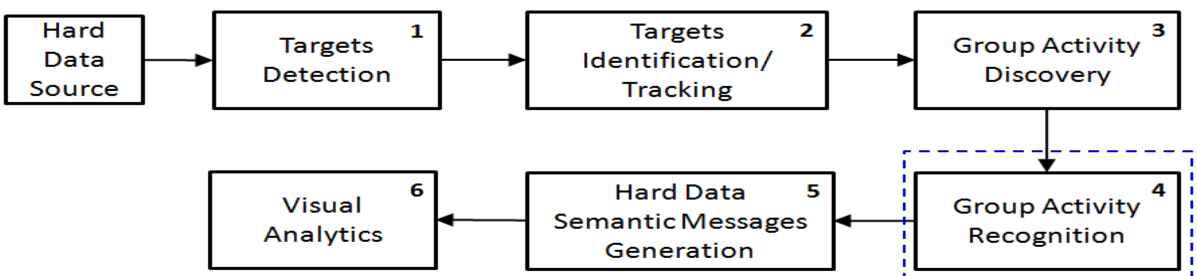


Figure 146: Conceptual Overview of Group Activity Monitoring System

Stage-1 is the target detection and refinement. This stage is typically involved with image processing techniques that facilitate elimination of noises, subtracting background, enhancing image via appropriate digital filters, and extracting features representing targets of interests. Stage-2 involves with spatiotemporal identification and tracking of targets using their trajectories motion analysis. Group activity discovery in stage-3 involves with the process of detecting presence of an activity performed by a group of entities. In Stage-3 characterization of detected GA is performed and their associated actions and activities are registered. In stage-4, correlation and association of detected GA actions and activities are verified and validated via the trained HMM's. For validation of group activities a set of pre-defined GA ontologies are employed. In stage-5, per each recognized group activity, an appropriate semantic message is generated based on a modified TML (Transducer Markup Language) data structure format. The composition of TML data structure format can be found in our reference [J.16] and not discussed here. The final stage is intended for Visual Analytics (VA) of multi-modality sensors and meant for further

analysis and inference of generated TML messages revealing comprehensive nature of different group activities [J.17][J.18].

3.4.7.2 HMM Model Architectures

HMM is a generative probabilistic model that can be trained to map observed sensory data onto latent hidden activity states. Hidden Markov Models can be considered as finite state machines where for each sequence of observation represents a state transition and an emission transition probability. One of the major goal of HMM is to determine the hidden state sequence ($x_1,x_2,\ldots x_t$) that corresponds to the observed sequence ($y_1,y_2\ldots y_t$) as shown in Figure 147. Three canonical problems have been addressed in HMMS namely, (1) Evaluate Probability of Sequence: given the model parameters, compute the probability of a particular output sequence. This is solved by the Forward or Backward algorithms [xx]. (2) Decoding: given the model parameters, find the most likely sequence of (hidden) states which could have generated a given output sequence. This is typically solved by the popular Viterbi algorithm in conjunction with posterior decoding (Maximum Likelihood (ML) or Maximum-a-posteriori (MAP));

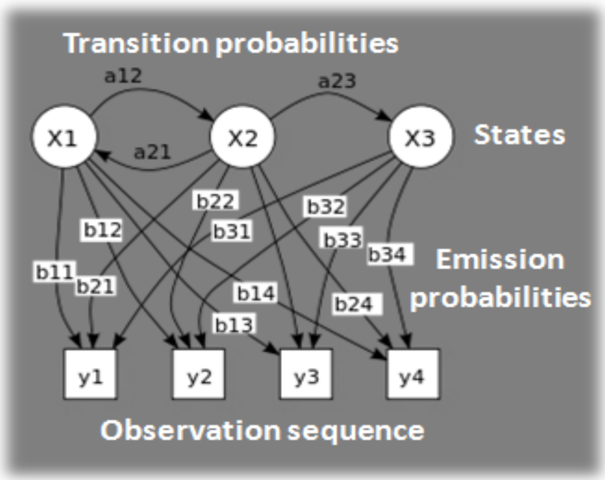


Figure 147: HMM Probabilistic Parameters

(3) Find the Model: given an output sequence, find the most likely set of state transitions and output probabilities. This problem has been also solved by the well-known Baum-Welch algorithm [xx]. Figure 148 shows a general structure of a HMM for predicting the most likelihood class of an input sequence.

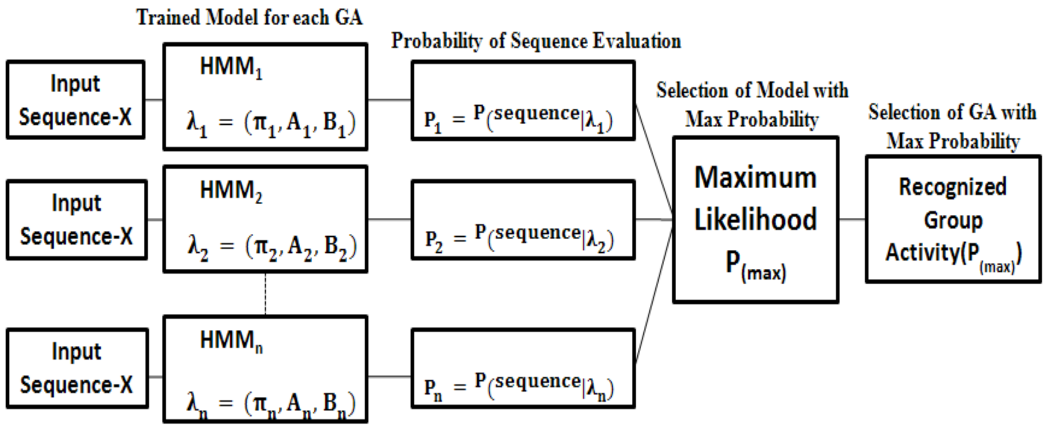


Figure 148: A HMM General Architecture for Recognition of Group Activities

Mathematically an HMM can be expressed as: $\lambda = (\pi, A, B)$, where A represents an (NxN) state transition matrix, B is an (NxM) observation probability matrix, and π is the initial state distribution. The activities (i.e., a sequence of recognized atomic actions) observed by the field sensors are coded and fed to a HMM for GA recognition. Training of HMM is achieved primarily by constructing some ontology representing known group activities patterns. Through this process a HMM recognizes the likelihood of a test input sequence representing an activity, and maps that to a known trained ontology class with certain likelihood scale. Given the complexity of input sequences in terms of order and content, different implication in interpreting the input sequence may arise. To overcome this issue, we introduce three different competing architectures of Hidden Markov Model: Cascaded HMM, Concatenated HMM and Context-based HMM. The following sections describe the pros and cons of each model.

3.4.7.3 HMMs Activity Modeling Design Concerns

Group Activity Recognition system is developed for predicting the group activities from the traced evidential sequential observations through modeling HMMs. Hidden Markov Models attempts to model dynamic systems whose latest output depends only on the current state of the system. Developed HMMs system infers the most likely sequence of states that can corresponds to a given input sequence. HMMs calculate the probability of a given sequence of outputs originated from the system. For each evidential observation, it is assumed that there exists a state transition and an emission transition. For developing a HMMs model, the following are the important parameters in developing HMMs which affect the likelihood measure of recognized activity: number of states and observations, number of iterations or the convergence factor, transition and emission probabilities and initial probability measures. On the other hand, generating the input sequence and modeling the activity prediction using the developed HMMs requires the consideration of different vital parameters. The following are the important parameters considered in generating the input sequence:

1) Number of events in a input sequence
2) Association and correlation between events
3) Spatiotemporal dependent events
4) Uncertainity involved in events recognition
5) Frequency of events/sequential bonded events

HMMs predict the type of activity involved based on the emerging sequential events. But when more events in a sequence are fed to HMMs, the chance of neglecting an occurred activity is more since HMMs are efficient in predicting the current involved activity (state). To avoid abandoning any activity being detected, the sequence is broken down to desired window frame i.e. required number of events in a sequence. The ideal number of events in a sliding window frame of sequence is the number of events existent in a trained HMMs ontology. Table 39 shows a sample of developed ontology for Group Activity Recognition. As seen in the ontology, three salient observations are used to construct each ontology pattern.

Table 39: Group Activity Ontology

Observation Sequence	Group Activity
Person_Carrying-Placed_in_Vehicle-Trunk_close	LOADING
Person_Picked-Person_Carrying-Placed_in_Vehicle	LOADING
Taken_from_Vehicle-Person_Carrying-Placed_in_Vehicle	LOADING
Person_Picked-Person_Carrying-Walking	UNLOADING
Trunk_open-Person_Carrying-Person_Dropped_object	UNLOADING
Trunk_open-Taken_from_Vehicle-Person_Carrying	UNLOADING
Taken_from_Vehicle-Person_Carrying-Left_Behind	ABNORMAL OBJECT DROPPING
Person_Carrying-Trunk_close-Left_Behind	ABNORMAL OBJECT DROPPING

To match the ontology using the HMMs, number of events in a window frame is chosen to be three. HMMs model predicts an activity for an input sequence. Three alternatives are identified in generating the input sequence to the HMMs model and are shown in Figure 149.

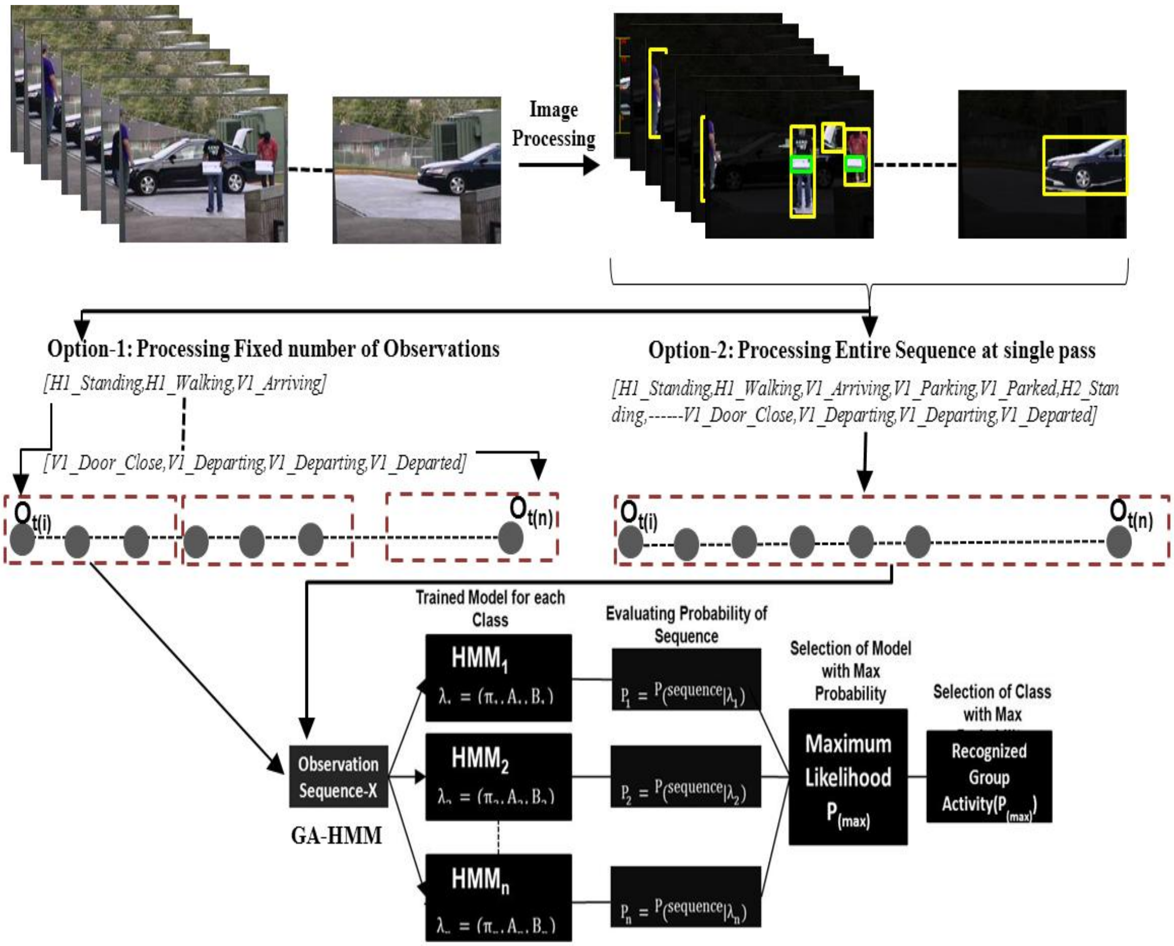


Figure 149: Activity Input Sequence

If *n* is the number of events in a given observed sequence, *w* is the desired number of events in a input sequence to the HMMs, then the number of input sequence is:

Case-1: $n \div w = 1$ since $n = w$

Case-2: $flooring(n \div w)$

Case-3: $(n - w + 1)$

In case-2, there may exist an instance where recent observations may be neglected. For example, if n is 11, w is 3, then the number of input sequence = 9. Two of the recent observations are neglected in input sequence generation. Also, some combinations of events in the sequence may not be in better correlation for predicting an activity. Whereas, in case-3, since sliding window is used to generate the sequences, all the events/observations are utilized in generating the input sequences. Both in case-2 and case-3, more than one activity is predicted since the number of input sequences are greater than one, provided that $(n > 2\times w)$ in case-2 and $(n > w)$ in case-3. A Maximum Likelihood (ML) is used for the predicted activities in case-2 and case-3 for identifying the most relevant activity.

When generating the input sequence, it is essential to identify the association and correlation between events. For example in a sequence (e_1-e_3-e_4), $e_1$and e_3 may not be associated or correlated with each other which could possibly end in predicting an incorrect activity. To perform the association and correlation between the events, equal weights are alloted to the ontology sequences. For example,consider an ontology (e_1-e_3-e_5)→Ontology-1, the associated weight for (e_1-e_3)= 0.5 and for (e_3-e_5) =0.5. A cumulative weights between two events are calculated for all the ontology used in the activity prediction.

The accumulated weight of all correlated pair in an ontology = 1. Therefore the correlated weight of each pair is given by $1 \div (number\ of\ events\ in\ an\ ontology - 1)$.

In our work, the number of events in each ontology are equal. The cumulative correlated weights of each pair in all ontology are calulcated through the following algorithm. Let 'x' be the correlated weight of each pair and 'y' be the number of times a pair occurred in all the ontology.

```
Int Correlation_wt(int y)
 {
   If (y=1) return x
   Else return Correlation_wt(y-1) + ((1-Correlation_wt(y-1))/2
 }
```

Correlation weights of each pair in the ontology is used in generating the input sequence. If the Correlation_wt is less than 'x', the combination of weight with the succeding event is calculated and the pair with the maximum weight is considered in the input sequence.

Percentage of confidence involved in identifying an event through pre-processing is also fed back to system. If an event is identified with less confidence, for example: event detected from a single source, its corresponding event index is removed and fed back to the model for tuning the activity prediction for better confidence thereby minimizing the uncertainity involved. The probability of an observation possibly occurring is computed by:

$$P(O(j)f(i+1))=P(O(j)f(i+1))+(1-P(O(j)f(i))/2$$

where O(j) is the observation of event j and P(O(j)f(i)) is the current probability of observation j occurring in frame f(i). Devloped modeling system accommodates to predict pertinent human activities by eliminating the events detected with less confidence. Removing trivial events and maintaining the distinct events in a input sequence also have significant impact in class (activity) prediction. Removing events of less importance i.e. trivial events helps in increasing the likelihood measures in appropriate HMM model selection. Spatiotemporal realtionship and frequency of events occurrence are also enforced in the activity detection. Certain operational activities can be recognized based on the frequency of occurrence of Sub Activities. Detecting the frequency of occurrence determines the intensity of the activity example: Situtaion awareness(S_A).

$$Frequency\ F(S\_A) > n\ \&\ Time\ elapse\ T < t, \qquad then\ Situation\ Awareness\ \uparrow$$

For eample if a person picks up an object, it is diagnosed as 'Object Removed' and if a person drops an object in a space-x, it is termed as 'Object Placed'. If 'Object Removed' and

‘Object Placed’ happens more frequent, then the activity is identified as ‘Loading’ and ‘Unloading’ operations.

The above discussed design parameters are effectively considered in generating the input sequence for each activity model. As mentioned earlier, three different model are used in activity prediction namely Cascaded model, Concatenated model and Context based model. The architecture of each model is discussed in the following sections.

3.4.7.4 Concatenated Modeling

The main goal of the concatenated modeling is to build a generative model that estimates the most likely label at each input sample (i.e. window frame sequence). In the developed HMMs model, each hidden state is directly associated with a specific label i.e. group activity to be detected. For demonstration purpose, the number of events in an input window sequence is considered as three. The output labels recognized from each window sequence are fused together for further processing. The fusion strategy employed is discussed in the end of Section 3.4.3. Figure 150 shows an example of a Concatenated model.

3.4.7.5 Cascaded Modeling

In the concatenated modeling, the inputs from the previous HMM may be used as an input for the next HMM i.e. output state from one HMM is used an observation in input sequence for the next HMM. Figure 151 shows an example of a Cascaded model where output from each HMM serves as an observation input to successive HMM. The output from each HMM is mapped with ontology for a possible match in a observation sequence. If the state is not considered in an input sequence, then model would behave as a concatenated model.

3.4.7.6 Context Based Modeling

Though the identified observations are continuous i.e. sequential observations, the distributions of observations for generating the input sequences can be effectively modeled by set of discrete classes independently for each context dimension. The test sequence is spitted into different input sequences based on the association between the events and a desired context as shown in Figure 152. For example in a context of monitoring a staircase, events related to Human-Vehicle interactions are not relevant to be modeled in activity prediction. The generation of each context input sequence, would be based on the design parameters as discussed in Section 3.4.7.1.

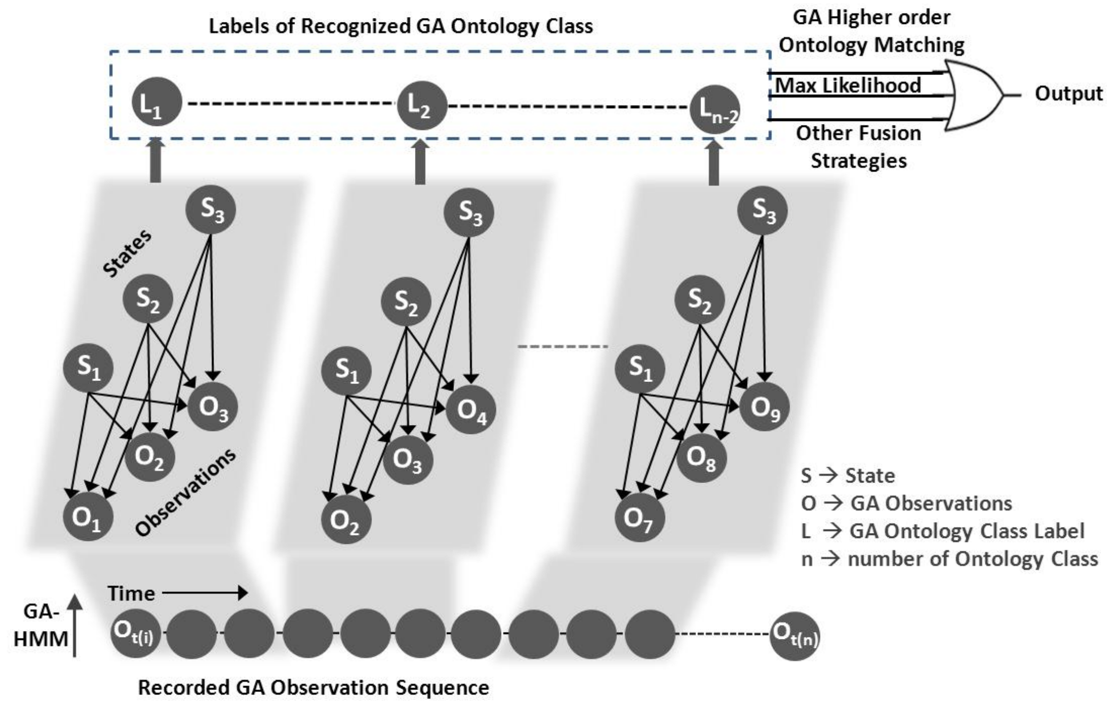


Figure 150: Concatenated HMMs

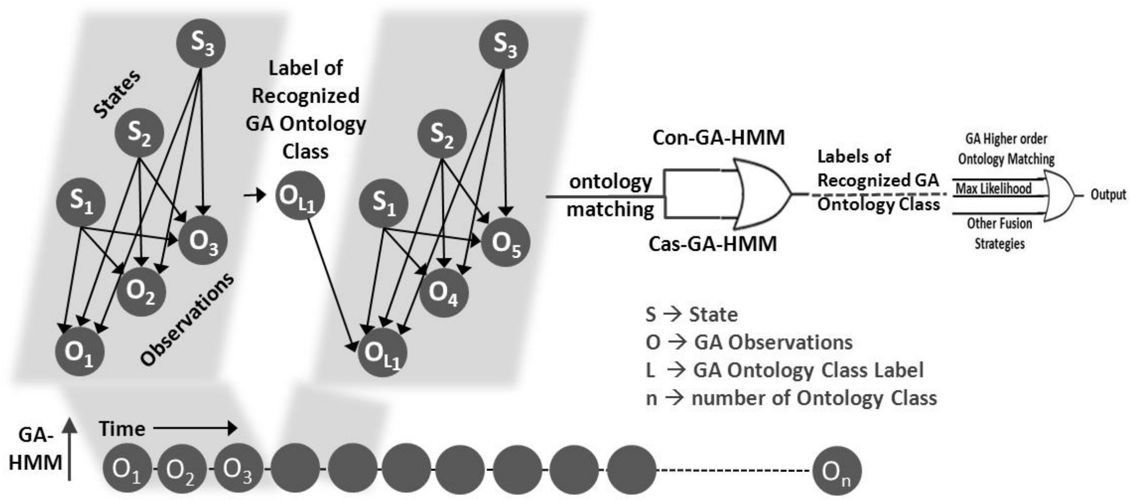


Figure 151: Cascaded GA-HMM (Cas-GA-HMM)

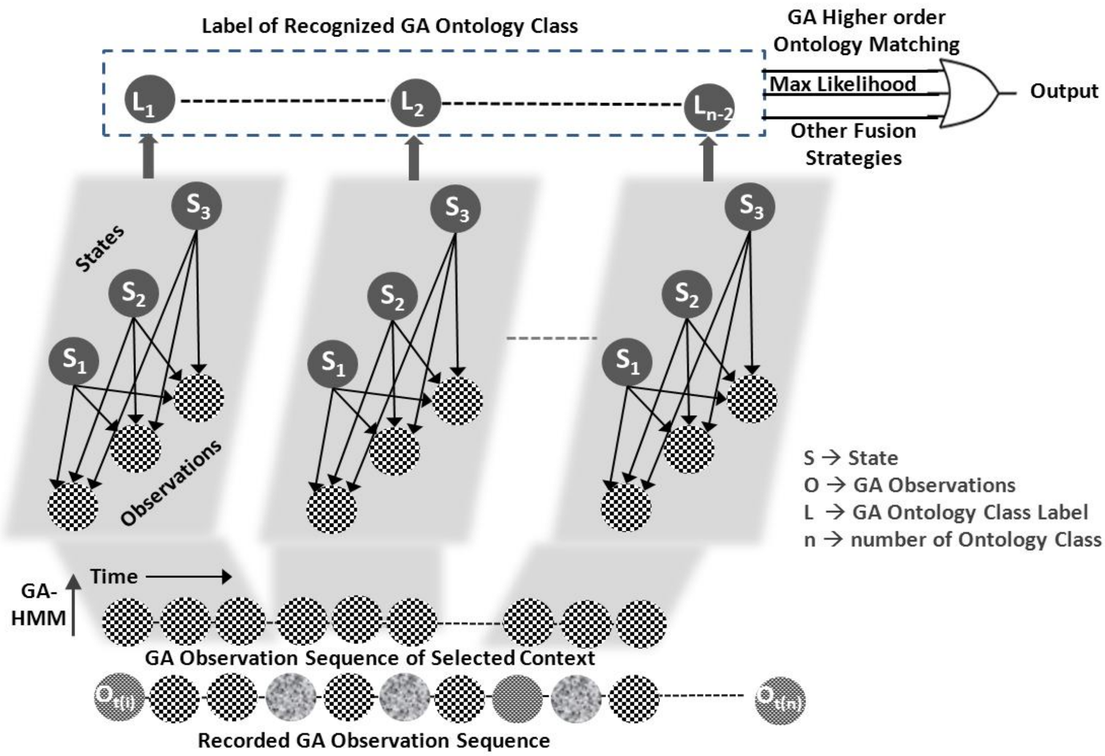


Figure 152: Hybrid Context based HMMs

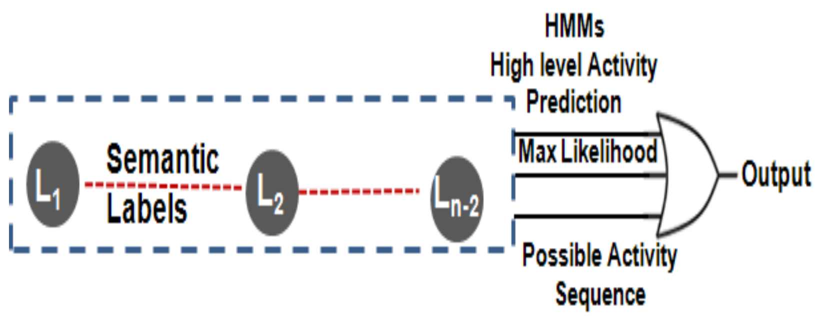


Figure 153: Processing of output semantic labels

Each model is selected based on the user requirements. The output semantic labels from the model are used for further processing. As shown in Figure 153, three possible processing are identified namely:

1. Feedback to Activity modeling for Higher order Activity prediction if matched with the high level ontology
2. Maximum Likelihood (ML) is applied to find out most likely occurred activity
3. Identifying the most possible activity sequence occurred in the scenario through matching the ontology.

The test sequence is split into different input sequences based on the association among the events and a desired context as shown in Figure 154. For example in a context of monitoring a staircase, events related to Human-Vehicle interactions are not relevant to predict activities relevant to a staircase. Therefore, depending on the context, an appropriate training ontology sequence would suffice to correct classify context-based group activities. The following reasons make CB-GA-HMM efficient and effective in recognizing GA in a scenario:

1. Minimizes the number of observations in an observation sequences to be modeled for GA recognition, thereby reducing the time complexity for GA recognition.
2. CB-GA-HMM can be used to focus on recognition of crucial type of group activities in an environment.
3. When context data is applied on GA ontologies during training, CB-GA-HMM is more effective since the testing can be applied with HMMs trained specifically with the GA ontology class of respective context as shown in Figure 154.

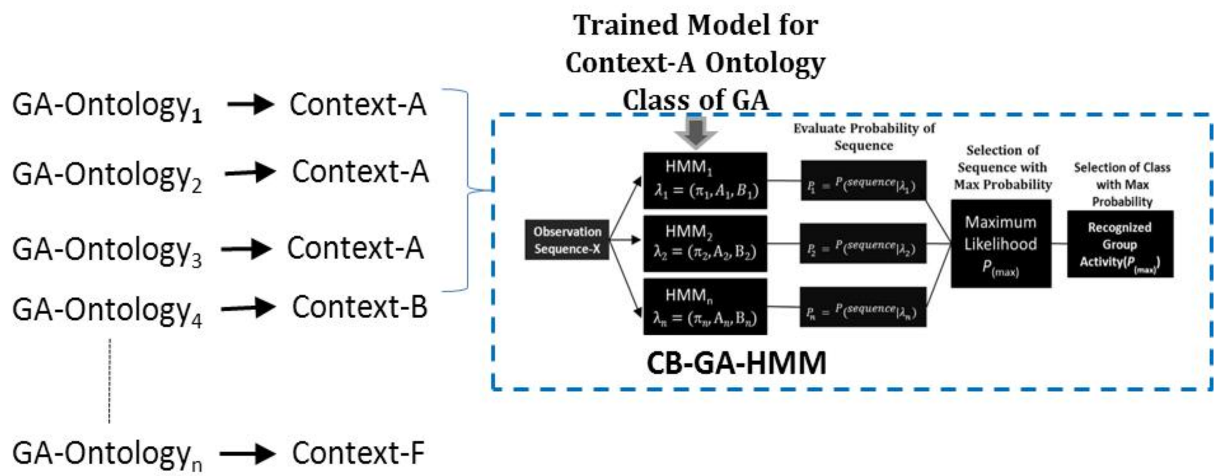


Figure 154: CB-GA-HMM for Context-A GA Ontology Class

3.4.7.7 State Transition of Object Handling:

Five different possible observations made in human handling objects (none, carrying small object, carrying large object, placing object, removing object), had been considered as shown in Figure 119. The size of object being used is subjective to target entity. State transitions of human object handling from one observation to other and the severity of each observation is shown below in Table 40.

Table 40: Object Handling State Transitions

	None	CSO	CLO	PO	RO
None	0.02	0.27	0.52	0.73	0.56
Carry small object (CSO)	0.18	0.25	0.62	0.48	0.61

Carry large object (CLO)	0.42	0.44	0.35	0.58	0.65
Placing object (PO)	0.40	0.60	0.62	0.48	0.58
Removing object (RO)	0.63	0.43	0.58	0.49	0.45

3.4.7.8 State Transition of Entity-Entity Relationship:

Relationships between two entities are analyzed based on observations as non-associative, approaching, united, cooperating and departing as depicted in Table 41.

Table 41: Entity-Entity Relationship State Transitions

	Non Associative	Approaching	United	Cooperating	Departing
Non Associative	0.08	0.50	0.62	0.75	0.40
Approaching	0.35	0.28	0.55	0.53	0.65
United	0.57	0.25	0.38	0.52	0.55
Cooperating	0.47	0.35	0.37	0.35	0.59
Departing	0.27	0.53	0.49	0.45	0.47

The chart above represents a template of state transition situational awareness severity 'probabilities' that two or more entities may render when their casual social interactions are observed through two sequential observations.

3.4.7.9 Visibility State Transitions:

Target visibility state modulation is done during target tracking and target behavioral pattern recognition. Suspiciousness here increases during each transition in successive observations. Three observations are deliberated such as normal, obscured and hidden. For example: given two sequential frames, subject-1 is in visible state at frame $t(x)$ and is hidden at frame $t(x+1)$ when an activity is recorded, then entropy of suspiciousness is believed to be 0.83. Weighed observations in visibility state transitions are shown in Table 42. Targets are assumed to be human or vehicle or a large object.

3.4.7.10 State Transition of Human Postures:

State transitions of human postures are studied in tracking of target's behavior. Human postures like standing, sitting, bending and laying down are considered from static image observations. Severity associated with sequential observations during the state transitions of human postural behavioral patterns is tabulated in the Table 43.

Table 42: Target Visibility State Transitions

VISIBILITY TRANSITION STATE			
	Normal	Obscured	Hidden
Normal	0.08	0.48	0.83
Obscured	0.37	0.35	0.79
Hidden	0.58	0.63	0.63

Table 43: Human Posture State Transitions

POSTURE STATE TRANSITION				
	Standing	Sitting	Bending	Laying
Standing	0.07	0.20	0.40	0.73
Sitting	0.23	0.17	0.43	0.65
Bending	0.23	0.38	0.28	0.63
Laying	0.52	0.48	0.53	0.58

3.4.7.11 State Transition of Human Kinematics:

Five behaviors such as motionless, walking, running, crawling and jumping are addressed for human kinematics state transitions analysis. This transition analysis is done on the sequential observations recorded through efficient semantic annotations. State transitions and severity colligated from sequential information from generated semantic message are shown in Table 44.

Table 44: Human Kinematics State Transitions

KINEMATIC STATE TRANSITION					
	Motionless	Walking	Running	Crawling	Jumping
Motionless	0.00	0.26	0.54	0.75	0.68
Walking	0.23	0.14	0.52	0.76	0.73

Running	0.42	0.41	0.43	0.76	0.73
Crawling	0.58	0.53	0.70	0.54	0.75
Jumping	0.53	0.41	0.63	0.79	0.48

3.4.7.12 State Transition of Distance to Target:

State transition associated with two images with respect to the distance between the entities to the target is identified using Table 45. By identifying the distance to the target, the entity's kinematics can also be identified through logical reasoning. For example: in frame t(1) subject-1 found far away the border of target and in frame t(2) subject-1 found inside the zone can reason that subject-1 moved fast towards the target.

Table 45: Distance-to-Target State Transitions

DISTANCE WITH RESPECT TO ZONE/TARGET				
	Far Away	Near Border	Ins Zone	Nearby Tar
Far Away	0.13	0.37	0.62	0.79
Near Border	0.61	0.30	0.59	0.75
Inside Zone	0.66	0.47	0.57	0.78
Nearby Target	0.70	0.53	0.58	0.75

For a smooth 'Distance to Target' state transition, a human need to follow the transition stated below:

Far Away → Near Border → Inside Zone → Nearby Target.

By normalizing and aggregating the above state transitions with respect to the context of atomic event categoriess, the model can predict severity of a given situation as human activities take place. Such estimation can further used as inference to generating appropriate semantic labels communication from within a surveillance system to a command control in form of tagged linguistic messages. This approach will improve situation awareness in persistent surveillance systems.

3.4.7.13 Proposed Semantic Message Generation

To fuse the heterogeneous information gathered from the acoustic and imagery data at the event detection level, a common TML structure for both sensor modalities was developed. The developed TML structure in a modified version of Transducer Mark-up Language data structure, originally introduced by OpenGIS research forum for communicating GIS sensors data, The modified version of TML introduced in this work, provide additional information that facilitates fusion of heterogeneous sensor data seamlessly. A TML message is, indeed, metadata associated with the sensor observation and contains attributes which describe the nature of detection the sensor has performed upon receiving an observation of the environment. Moreover, TML enables to transfer data and metadata at one single vector and enables to develop common strategy for data processing, multimodality fusion and visual analytics. The data contained in the TML encompass the AOI discussed in the Table 40. Messages generated will follow a short version of TML format for generating semantic scene labels. TML format used for generating messages is given below:

<data_ref="source">month/day/year,Thh:min:sec,message-id,Global_Space, Source_attribute,Local_Space,Reference_Space,Detection_Focus,ID,Attribute-1, Attribute-2,Attribute-3,Attribute-4,Attribute-5,Attribute-6,Group-ID,Confidence </dat_ref ="source">

Figure 155 illustrates the TML format used for generating messages describing a group activity.

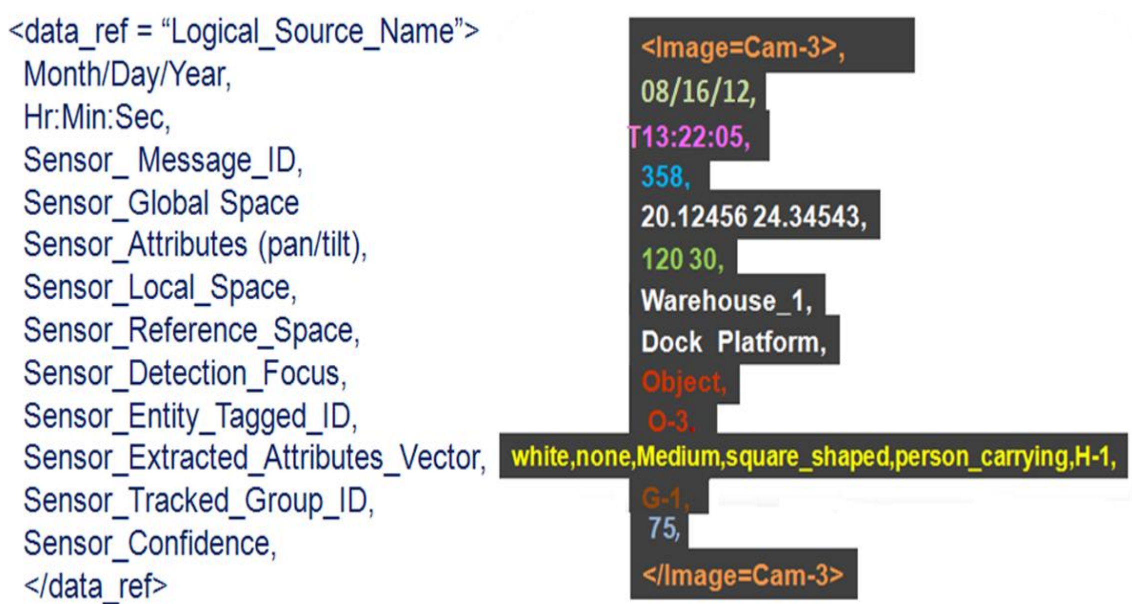


Figure 155: TML Format for Group Activity Characterization

This TML format accommodates the time and location of the data source. GPS coordinates specifies the location of the data source (e.g. surveillance camera) and is denoted as *Global_Space*. *Local_Space* denotes the environment location (eg: Warehouse, parking lots, and

etc.) and the R*eference_Space* denotes the location of the target's reference (eg: dock platform, vehicle hood zone, and etc.). *Detection_Focus* specifies the target entity or group activity under focus as shown in the Table 32. *Entity_Tagged_ID* denotes the unique ID tagged for a target entity. The six attributes of the detection focus denotes the GA characteristics discussed in the earlier section and the Table 40. *Group-ID* denotes the belonging of the target to the specified group. It also helps in tracking a target entity and his/her GA interactions with members of same group or members of other groups. A group in an activity is identified by the association of the individuals to the other members in a group. For example, if a set of people exiting from a vehicle, each individual can be considered as part of a group. Group can also be identified when a set of people grouped together in performing certain goal oriented actions such as group walking, group talking, group running, group carrying objects in performing a loading/unloading of objects activity. Confidence specifies the percentage of confidence in detecting the *Detection_Focus*.

Each messages generated carries a detection confidence. Detection confidence can be specified in different ways namely 1) an average of confidence level of different attributes (Avgconf) 2) average of confidence level of attributes multiplied with a respective weight factor (Avgconf_wt) 3) max or min of the attribute confidence (Confmax) or (Confmin) 4) average of confidence after a threshold limit for each attribute (Avgthrs_conf), etc., Let Conf = {Conf1, Conf2,…… Confn }, then:

$$\mathrm{Avg_{conf}} = (\mathrm{Conf_1} + \cdots \mathrm{Conf_n})/\mathrm{n}$$

$$\mathrm{Avg_{conf\_wt}} = \frac{\sum_{i=1}^{n} \mathrm{W_i Conf_i}}{\sum_{i=1}^{n} \mathrm{W_i}}$$

$$\mathrm{Avg_{conf\_wt\_thrs}} = \frac{\sum_{i=1}^{n} \mathrm{W_i Conf_i}}{\sum_{i=1}^{n} \mathrm{W_i}} \text{ if } \mathrm{W_i} > \mathit{threshold\_value}$$

In image processing, the detection confidence can be increased or decreased depending on the number of frames being processed. In this work, an average confidence from the detected sensor extracted attributes is calculated for each TML message as shown in the last relationship above.

The fused TML messages are stored in the TML pipeline for further processing based on the required visual analytics. The generated TML messages can be generated in two different forms: a) TML message in semantic annotations (i.e. textual form) and b) TML in numerical coded (i.e. each semantic annotation is numerically indexed). The TML format used in this research for describing the experimental scenarios are in the textual form. A sample of TML messages generated from a GA scenario is shown below:

<Image=Cam-5>01/27/10,T12:0:2,2,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,None,Black,Sedan,None,None,Arriving,None,None,50</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:3,3,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,None,Black,Sedan,Normal,North,Arriving,None,None,75</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:4,4,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,V-1,Black,Sedan,Normal,North,Arriving,None,None,90</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:5,5,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None,Vehicle,V-1,Black,Sedan,Normal,North,Parking,None,None,75</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:6,6,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,V-1,Black,Sedan,Normal,North,Parking,None,None,87.5</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:7,7,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,V-1,Black,Sedan,Normal,North,Parking,Door_Open,None ,75</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:8,8,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Vehicle,V-1,Black,Sedan,Normal,North,Parking,Door_Open,None,87.5</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:8,8,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Human,H-1,Grey,Standing,None,V-1,None,Passenger,None,75</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:9,9,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Human,H-1,Grey,Standing,None,V-1,Distant_Interaction,Passenger,None,87.5 </Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:12,12,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Human,None,White,Standing,None,None,None,None,None,50</Image=Cam-5>

<Image=Cam-5>01/27/10,T12:0:15,15,xx.xxxxxx yy.yyyyyy,90 30,Warehouse-1,None, Human,H-3,Black,Standing,None,V-1,None,Passenger,G-1,75</Image=Cam-5>

Note: GPS coordinates in the above TML messages is concealed and noted as (xx.xxxxxx yy.yyyyyy).

3.4.7.14 Imagery Annotation Results Analysis

This section describes an experiment carried out for validation of group activity prediction in a context of human-vehicle, human-human and human-object interactions in a monitored environment. The tested 'Loading of Objects' scenario is as shown in Figure 145 is as follows: Human-1 waits for the group in the Safe House; Vehicle-1 arrives and parks near the Safe House; Human-2 (Group Leader) and other group members Human-3, Human-4 exits the vehicle and group members unloads objects from the vehicle to the Safe House; Human-2 and Human-1 interacts with each other; Human-2,Human-3 and Human-4 enters the vehicle and departs.

The recorded observations from this scenario are:

[*H1_Standing, H1_Standing, H1_Standing, H1_Walking, V1_Arriving, V1_Arriving ,V1_Parking, V1_Parking, V1_Parked, H2_Standing, H2_Moving, H1_Moving, H2_Moving, H1_Moving, V1_Door_Open, H3_Standing, H4_Standing, Group_Merging, V1_trunk_Open, H3_O1_Taken_from_Vehicle, H4_O2_Taken_from_Vehicle, H3_O3_Taken_from_Vehicle, H4_O4_Taken_from_Vehicle, H3_O4_Person_Carrying, H3_O4_Person_Carrying, H1_O5_Taken_from_Vehicle, H2_Moving, Distant_Interaction, V1_Trunk_Close, H1_O6_Person_carrying, H1_O6_Person_carrying, V1_Door_Open, V1_Door_Open, V1_Trunk_Open, V1_Door_Open,V1_Trunk_Close, Distant_Interaction, V1_Door_Open, V1_Door_Close, V1_Door_Close, V1_Departing, V1_Departing, V1_Departed*]

These observations are fed to the developed HMMs for predicting the type of GA taken place in the safe house.

3.4.7.15 Imagery Data Processing and Interface

Figure 156 and Figure 157 presents two screen shots for the newly developed interface for Imagery Data Processing and Annotation.

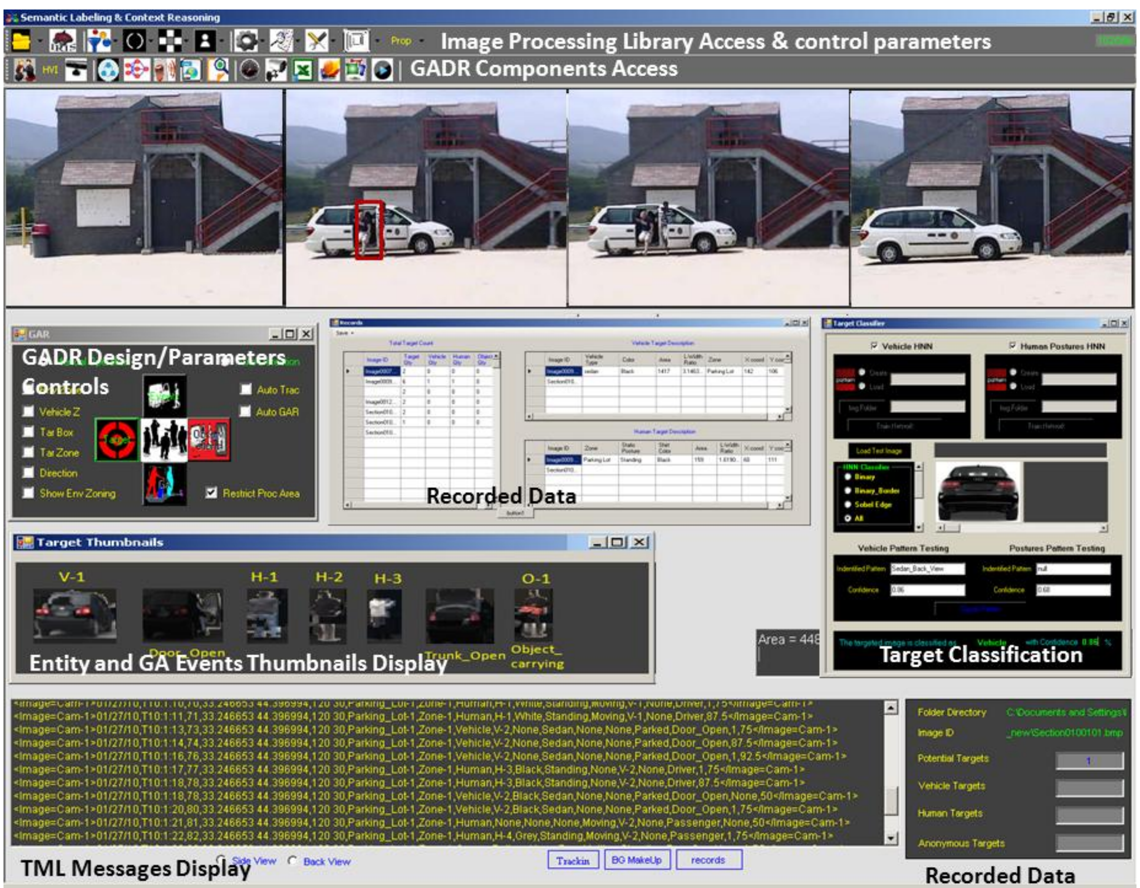


Figure 156: A Screen Shot of Interface Developed for Visual Imagery Processing and Annotation

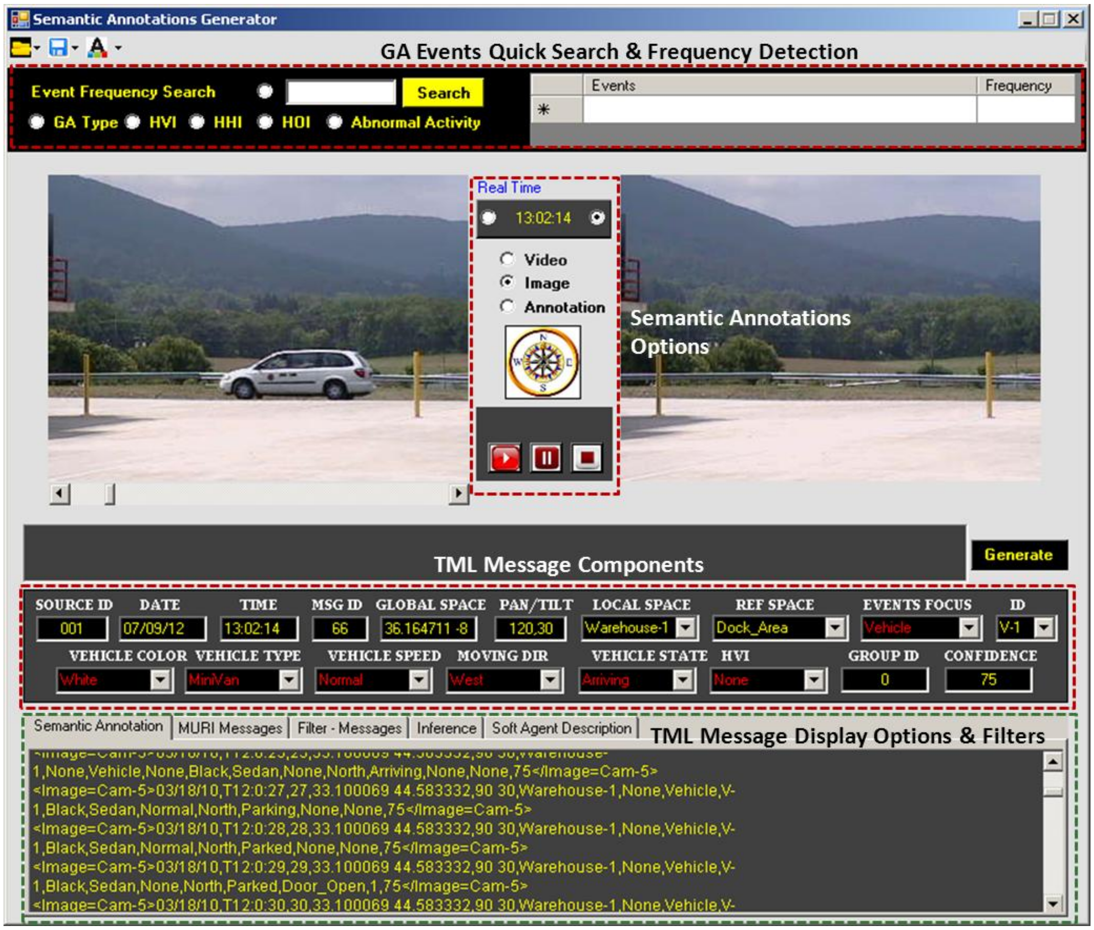


Figure 157: Another Screenshot of the Interface for Image Processing and Annotation

3.4.8 TSU References

[J.1] Elangovan, V., and Shirkhodaie, A., "A Robust Technique for Group Activities Recognition Based on Fusion of Extracted Features in Video Streams", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.

[J.2] Elangovan, V., Bashir, A., and Shirkhodaie, A., "A Multi-attribute Based Methodology for Vehicle Detection & Identification", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.

[J.3] Elangovan, V., Alkilani, A., and Shirkhodaie, A., "A Multi-Modality Attributes Unmasking Scheme for Group Activity Characterization and Data Fusion", IEEE Intelligence and Security Informatics (ISI), Seattle, WA, June 2013.

[J.4] Alkilani, A., and Shirkhodaie, A., "Acoustic Recognition of Human-Object Interactions in Persistent Surveillance Systems", SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.

[J.5] Shirkhodaie, A., Elangovan, V., Habibi, M. S., and Alkilani, A., "A Decision Support System for Fusion of Soft and Hard-sensor Information Based on Latent Semantic

Analysis Technique”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.

[J.6] Habibi, M. S., and Shirkhodaie, A., “Mining Patterns in Persistent Surveillance Systems with Smart Query and Visual Analytics”, SPIE Defense, Security and Sensing Conference, Baltimore, MD, April 2013.

[J.7] Shirkhodaie, A., Rababaah, H., "Multi-Layered Context Impact Modulation for Enhanced Focus of Attention of Situation Awareness in Persistent Surveillance Systems", Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, Proceedings of SPIE Vol. 7710 (SPIE, Bellingham, WA 2010), paper number 7710.

[J.8] Shirkhodaie, A., Rababaah, H., "Multi-layered context impact modulation for enhanced focus of attention of situational awareness in persistent surveillance Systems," in Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2010, Proceedings of SPIE Vol. 7710, 2010.

[J.9] Rababaah, H., Shirkhodaie, A., “Semantic Labeling of Non-stationary Acoustic Events in Persistent Surveillance System,” in the SPIE Photonic West, January 22-27, 2011, San Francisco, CA.

[J.10] Elangovan V., and Shirkhodaie, A., “A Survey of Imagery Techniques for Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,“ in the SPIE Defense and Security Conference, Orlando, FL, 2011.

[J.11] Shirkhodaie, A., Rababaah, A., Elangovan V., “Acoustic and Imagery Semantic Labeling and Fusion of Human-Vehicle Interactions,” in the SPIE Defense and Security Conference, Orlando, FL, 2011.

[J.12] Elangovan V., and Shirkhodaie, A., “Context-Based Semantic Labeling of Human-Vehicle Interactions in Persistent Surveillance Systems,” in the SPIE Defense and Security Conference, Orlando, FL, 2011.

[J.13] Shirkhodaie, A., “Perceptual Semantic Labeling of Human-Vehicle Interactions (HVI)”, Second Annual Human and Light Vehicle Detection Workshop, Baltimore, MD, May, 2011.

[J.14] Shirkhodaie, A., “Semantic Labeling of Human-Vehicle Interactions Via Acoustic Events Characterization and Inference,“ Second Annual Human and Light Vehicle Detection Workshop, Baltimore, MD, May, 2011.

[J.15] Elangovan ,V., Shirkhodaie, A., “Human Activity Discovery & Recognition based on State Transitions Modeling in Persistent Surveillance Systems,” the SPIE 2012 Electronic Imaging: Science and Technology, Algorithms and Systems VI, 22 - 26 January 2012, San Francisco, CA.

[J.16] Vinayak Elangovan, Amir Shirkhodaie, “Team activity analysis and recognition based on Kinect depth map and optical imagery techniques“, SPIE 2012 Defense,

Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, Paper number 8392-30, April 23, 2012, Baltimore, MD.

[J.17] Vinayak Elangovan, Amir Shirkhodaie, "Team activity analysis and recognition based on Kinect depth map and optical imagery techniques", SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, Paper number 8392-30, April 23, 2012, Baltimore, MD.

[J.18] Vinayak Elangovan, Amir Shirkhodaie, "Recognition of Human Activity Characteristics Based on State Transitions Modeling Technique", SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, Paper number 8392-43, April 23, 2012, Baltimore, MD.

[J.19] Amjad Alkilani, and Amir Shirkhodaie, "A survey on acoustic signature recognition and classification techniques for persistent surveillance systems ", SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, Paper number 8392-28, April 23, 2012, Baltimore, MD.

[J.20] Mohammad Habibi, and Amir Shirkhodaie, "A Survey of Visual Analytics for Knowledge Discovery and Content Analysis," SPIE 2012 Defense, Security, and Sensing Symposium, Conference: Signal Processing, Sensor Fusion, and Target Recognition XXI, Paper number 8392-27, April 23, 2012, Baltimore, MD.

[J.21] Amir Shirkhodaie, "Perceptual Semantic Labeling of Human-Vehicle Interactions (HVI)," 2nd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 3-4, 2011, Baltimore, MD.

[J.22] Amir Shirkhodaie, and Amjad Alkilani, "Acoustic Sounds Recognition and Classification for Group Activities Discovery," 3rd Annual Human, Light Vehicle and Tunnel Detection Workshop, May 16-17, 2012, Mississippi.

[J.23] P.A. Viola and M.J. Jones. "Rapid Object Detection Using a Boosted Cascade of Simple Features. In CVPR (1), pp. 511-518, 2001.

[J.24] P. Simard, L. Bottou, P. Haner, and Y. L. Boxlets, "A Fast Convolution Algorithm for Signal Processing and Neural Networks. In NIPS, 1998.

[J.25] Herbert Bay , Andreas E., Tinne T. , and Luc Van G., "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008.

3.5 Iona College

List of papers submitted or published:

- Papers Published in Peer-reviewed Journals
  - Yager, R. R., "Pythagorean membership grades in multi-criteria decision making," IEEE Transaction on Fuzzy Systems 22, 958-965, 2014.
  - Yager, R. R. and Alajlan N., "Probability weighted means as surrogates for stochastic dominance in decision making," Knowledge-Based Systems 66, 92-98, 2014.

- Yager, R. R. and Alajlan, N., "Probabilistically weighted OWA aggregation," IEEE Transactions on Fuzzy Systems 22, 46-56, 2014.
- Yager, R. R. and Alajlan, N., "On characterizing features of OWA aggregation operators," Fuzzy Optimization and Decision Making 13, 1-32, 2014.
- Yager, R. R. and Alajlan N., "A note on mean absolute deviation," Information Sciences *279,* 632-641, 2014.
- Yager, R. R., "Joint cumulative distribution functions for Dempster–Shafer belief structures using copulas," Fuzzy Optimization and Decision Making 12, 393-414, 2013.
- Yager, R. R. and Alajlan, N., "Decision making with ordinal payoffs under Dempster–Shafer type uncertainty," International Journal of Intelligent Systems 28, 1039-1053, 2013.
- Yager, R. R., "On forming joint variables in computing with words," International Journal of General Systems 42, 820-837, 2013.
- Yager, R. R. and Abbasov, A. M., "Pythagorean membership grades, complex numbers and decision-making," International Journal of Intelligent Systems 28, 436-452, 2013.
- Yager, R. R., "Membership modification and level sets," Soft Computing 17, 391-399, 2013.
- Yager, R. R. and Alajlan, N., "Measure based representation of uncertain information," Fuzzy Optimization and Decision Making 11, 363-385, 2012.
- Yager, R. R., "On prioritized multiple criteria aggregation," IEEE Transactions on Systems, Man and Cybernetics: Part B 42, 1297-1305, 2012.
- Yager, R. R., "Participatory learning of propositional knowledge," IEEE Transactions on Fuzzy Systems 20, 715-727, 2012.
- Yager, R. R., "Measures of assurance and opportunity in modeling uncertain information," International Journal of Intelligent Systems 27, 776-796, 2012.
- Yager, R. R., "Entailment principle for measure-based uncertainty," IEEE Transactions on Fuzzy Systems 20, 526-535, 2012.
- Yager, R. R., "Dempster-Shafer structures with general measures," International Journal of General Systems 41, 395-408, 2012.
- Yager, R. R., "Conditional approach to possibility-probability fusion," IEEE Transactions on Fuzzy Systems 20, 46-56, 2012.
- Yager, R. R., "On Z-valuations using Zadeh's Z-numbers," International Journal of Intelligent Systems 27, 259-278, 2012.
- Yager, R. R., "Expansible measures of specificity," International Journal of General Systems 41, 247-263, 2012.
- Yager, R. R., "Set measure directed multi-source information fusion," IEEE Transactions on Fuzzy Systems 19, 1031-1039, 2011.
- Yager, R. R., "Firing fuzzy rules with measure type inputs," IEEE Transactions on Fuzzy Systems, (To Appear).

- Yager, R. R., "Stochastic dominance for measure based uncertain decision making," International Journal of Intelligent Systems, (To Appear).
- Yager, R. R. and Mesiar R., "On the transformation of fuzzy measures to the power set and its role in determining the measure of a measure," IEEE Transactions on Fuzzy Systems, (To Appear).
- Yager, R. R., "An intuitionistic view of the Dempster-Shafer belief structure," Soft Computing (To Appear).
- Elmore, P., Petry, F. E. and Yager, R. R., "Comparative measures of aggregated uncertainty representations," Journal of Ambient Intelligence and Humanized Computing, (To Appear).
- Yager, R. R., "A fuzzy measure approach to systems reliability modeling," IEEE Transactions on Fuzzy systems, (To Appear).

- Papers published in peer-reviewed conference proceedings
  1. Yager, R. R., "Measure inputs to fuzzy rules," Information Processing and Management of Uncertainty in Knowledge-Based Systems: Proceedings of the 15th IPMU International Conference, Part I, Montpellier, France, Edited by A. Laurent, O. Strauss, B. Bouchon-Meunier and R. Yager, 577-587, 2014.
  2. Yager, R. R. and Yager, R. L., "Soft retrieval and uncertain databases," 47th Hawaii International Conference on Sytems Science, Waikola, HI, 934-942, 2014.
  3. Yager, R. R., "Pythagorean fuzzy subsets," Proceedings of the Joint IFSA Congress and NAFIPS Meeting, Edmonton, Canada, 57-61, 2013.
  4. Yager, R. R. and Petry, F. E., "Intuitive decision-making using hyper similarity matching," Proceedings of the Joint IFSA Congress and NAFIPS Meeting, Edmonton, Canada, 386-389, 2013.
  5. Yager, R. R., "On a view of Zadeh's Z-numbers," Advances in Computational Intelligence-Procceedings of the 14th International Conference on Information Processing and of Uncertanity in Knowledge-Based Systems (IPMU) Part 3, Catania, Italy, Springer:Berlin, 90-101, 2012.

- Other presentations
  1 Plenary Lecture Tenth International Conference on Hybrid Intelligent Systems 2010: On the Fusion of Soft Information

  2 Plenary Lecture at IEEE 2010 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2010): Intelligent Learning Methods for Soft Computing Systems

  3 Plenary Lecture North American Fuzzy Information Processing Society NAFIPS 2011: A Measure Based Approach to the Fusion of Uncertain Information

  4 Plenary Lecture North American Fuzzy Information Processing Society NAFIPS 2012: Fuzzy Methods for Constructing Multi-Criteria Decision Functions

5 Plenary Lecture- Second World Conference on Soft Computing 2012: Social Network Computing

6 Plenary Lecture Third World Conference on Soft Computing 2013 Berkeley Intelligent Aggregation Methods for Decision Making and Learning

7 Plenary Fourth World Conference on Soft Computing 2014: Information Representation and Fusion for Decision Making

8 Invited Talk U.S. Army Research Laboratory, Aberdeen Proving Grounds, 2013: Issues in Hard Soft Information Fusion

l) Manuscripts

1 Yager, R. R., "Combining various types of belief structures," Technical Report MII-3407 Machine Intelligence Institute, Iona College, New Rochelle, NY, 2014.

2 Yager, R. R. and Alajlan N., "Evaluating belief structure satisfaction to uncertain target values " Technical Report MII-3415 Machine Intelligence Institute, Iona College, New Rochelle, NY, 2014.

m) Books and Book Chapters

1 Yager, R. R., "Intelligent aggregation and time series smoothing," In Time Series Analysis, Modeling and Applications, Pedrycz, W. and Chen, S. M. (Eds), Springer, Heidelberg, 53-75, 2013.

2 Yager, R. R., Reformat M. Z., and Alajlan N., Intelligent Methods for Cyber Warfare. Heidelberg: Springer, (To Appear).

3 Zadeh, L. A., Abbasov, A. M., Yager, R. R., Shahbazova, S. N. and Reformat, M. Z., Recent Developments and New Directions in Soft Computing, Springer: Heidelberg, 2014.

4 Laurant, A., Strauss O., Bouchon-Meunier B., and Yager R. R., Information Processing and Management of Uncertainty in Knowledge-Based Systems-Proceedings 15th International Conference IPMU Parts 1-3. Berlin: Springer, 2014.

5 Masulli, F., Pasi, G. and Yager, R. R., Fuzzy Logic and Applications, Lecture Notes in Computer Science 8256, Springer: Heidelberg, 2013.

6 Yager, R. R., Abbasov, A. M., Reformat, M. Z. and Shahbazova, S. N., Soft Computing: State of the Art Theory and Novel Applications, Springer: Heidelberg, 2013.

7 Greco, S., Bouchon-Meunier, B., Coletti, G., Fedrizzi, M., Matarazzo, B. and Yager, R. R., Advances in Computational Intelligence- Procceedings of the 14th International Conference on Information Processing and of Uncertanity in Knowledge-Based Systems (IPMU) Parts 1-4, Catania, Italy, Springer: Berlin, 2012.

Honors and Awards –

* Honorary doctorate degree *honoris causa* State University of Library Studies and Information Technologies, Sofia Bulgaria. Sept 6, 2012

* Honorary doctorate degree *honoris causa* Azerbaijan Technical University, Baku Azerbaijan, December 3, 2012

* Honorary Committee: The Second International Fuzzy Systems Symposium, Ankara, Turkey, 2011.

* Co-author with Fred Petry of the Naval Research Laboratory at Stennis Space Center of NRL Berman ARPAD Award, Annual Research Publications Award established to recognize the authors of the best NRL publications each year 2011

* Honorary Professor: Aalborg University Faculty Engineering and Sciences Denmark, 2010-13

* Honorary Chair the 2013 IEEE International Conference on Granular Conputing, Beijing

* Honorary Co-chair: the 7th IEEE Intelligent Systems'2014 Conference in Warsaw, Poland:

* Lifetime Outstanding Achievement Award International Fuzzy Systems Association 2011

* Ebrahim Mamdani Award for Outstanding Contributions in Fuzzy Computation presented by BRICS-CCI & CBIC Committee September 2013 Porto de Galinhas Brazil

* Recipient of the The Irene Hammerbacher Outstanding Faculty Research Award from Iona College 4/14

Titles of Patents disclosed during the reporting period –

Patents awarded during the reporting period –

Faculty

Name	Percent Supported
Ronald Yager	17%
Total Number:	1

Technology transfer

n) We have worked closely with researchers at the Stennis Space Center.

o) We have cooperated with Caterpillar Inc on a project using our technologies

p) We collaborated with Ford Motor Company researchers.

3.5.1 Uncertainty Modeling

The problem of multi-source information fusion arises in many disciplines. Here we have multiple pieces of information about some variable that we are interested in combining to get a unified view of our knowledge of the value of the variable. Our interest has been focused on the problem of "hard/soft" information fusion. In this environment the hard information is typically provided by some electro–mechanical sensor while the soft information is generally of a linguistic type provided by some human or obtained from the Internet. One fundamental issue here is the different nature of these types of information. Typically the sensor-based information is expressed in terms of a probability distribution. The soft information, while usually more difficult to formulate, can often be expressed using a fuzzy set and an associated possibility distribution. The different modes for representing these two types information and perhaps other types of information makes the problem of fusion a difficult task.

During this research we looked in detail at the multi-mode fusion problem and looked at different approaches to accomplish this task. Among those we looked at was a general framework for the fusion of different types of uncertain information based upon the use of a type of monotonic set measure called a fuzzy measure. While this approach provides the potential for the fusion of a large variety of different formulations of uncertainty we particularly investigated its potential for the fusion of probabilistic and possibilistic information. We showed that the fusion of these two types of uncertainty can result in an uncertainty formulation of a new type. We studied the properties of this type of uncertainty model.

As an alternative method for hard soft information fusion we developed a framework for fusing a probabilistic and possibilistic uncertainty based on the idea of conditioning the probability distribution with respect to the possibility distribution. We first investigated the problem of fusing multiple possibility distributions. In this case we particularly looked at the issue of normalization and suggest a generalized approach to the normalization of conflicting possibility distributions. We then look at the fusion of multiple probability distributions. Finally we suggested fusing the probabilistic and possibilistic uncertainty based on the idea of conditioning the probability distribution with respect to the possibility distribution. This approach results in a probability distribution as the fused information.

In the multi-source fusion problem, in addition to having a collection of pieces of information that must be fused, we need to have some expert provided instructions on how to fuse these pieces of information. These instructions can be provided by a human expert or by a machine learning system. Generally these instructions can involve a combination of linguistically and mathematically expressed directions. We considered the fundamental task of how to translate these instructions into formal operations that can be applied to our information. This task requires the use of aggregation operators. With this in mind we developed a framework for the aggregation of the types of monotonic set measures we used to model our uncertain information.

Recent interest has focused on imprecise uncertainty measures, the so-called second order uncertainty. In the case of the probability measure this corresponds to situation in which rather than knowing the precise probability of an event we only know an interval in which the probability lies. The Dempster-Shafer belief structure provides a framework for the representation of a wide class of imprecise (second order) uncertainty measures. When using

these structures to represent imprecise uncertain information in the multi–source environment we are faced with the problem of fusing multiple Dempster–Shafer belief structures. The earliest prescribed approach for fusing Dempster-Shafer belief structures was the Dempster rule. Researchers have raised some concern about Dempster's use of normalization in addressing issues arising from the intersection of conflicting focal elements. This concern has initiated interest in providing alternatives to Dempster's rule. During our research we have developed some alternative approaches to the fusion of Dempster-Shafer belief structures that circumvents the problem of normalization.

An important task in the problem of hard/soft information fusion is the translation of linguistically expressed information into a formal representation. The fuzzy set based theory of Approximate Reasoning (AR) provides a technology for representing and manipulating human sourced linguistically expressed soft information. Many applications based on these ideas can be found in the literature. An important component of this framework is its ability to represent imprecise and uncertain information using the concept of a generalized constraint statement. A typical example of such a statement is Johns *is near the river*, where we are assigning the variable John's location the value near the river. In this statement uncertainty is associated with the term near-river, the value of the variable. We extended the capability of this framework by considering an additional source of uncertainty in the generalized constraint statement, uncertainty with respect to variable itself. Formally the generalized constraint statement is of the form Attribute (Object) is Value. In the preceding the attribute is Location, the object is John and the value is near-river. We considered situations in which there exists some uncertainty with respect to object itself as exemplified by the statement ***a tall men*** *is near the river*. Here the attribute is still location and the value is still near-river but the object is an uncertain object, a tall men. We provided for representation of these doubly uncertain constraints and a mechanism for making inferences from these statements

In many situations the ultimate goal of information fusion is to aid in the solution of some multi–criteria decision problem, help in selecting an alternative that best satisfies a collection of criteria. Since the output of the hard/soft information fusion is often an imprecise uncertain value, we are interested in the issue of determining an alternatives satisfaction to a criterion when the alternative's associated attribute value is imprecise. We developed two approaches to the determination of criteria satisfaction in this uncertain environment, one based on the idea of containment and the other on the idea of possibility. We were particularly interested in the case in which the imprecise data is expressed in terms of a trapezoidal type possibility distribution. We provided an algorithmic solution to this problem enabling it to be efficiently implemented in a digital environment.

A common example of a human sourced observation is a statement of the type "I am pretty sure that the number of enemy soldiers is about 300." Here *pretty sure* is the confidence in the observation of "about 300 soldiers." Here we have some soft observation of the value of a variable V as well as some soft linguistic indication of our confidence (probability of the truth) associated with our observation. We investigated the concept of a Z-numbers recently introduced by Zadeh and studied its potential for representing the type of information described above. We showed how this kind of information can be formally modeled as a possibility distribution over probability distributions associated with V. We looked at the problem of the fusion of multiple statements of this kind as well as problems related inference, reasoning and decision with this type of information.

We worked further with the Dempster-Shafer belief structure and its role in the representation of uncertain information of both hard and soft types. We considered the problem of joining multiple Dempster-Shafer structures over different attributes. To this end we investigated the formation of the cumulative distribution function (CDF) for these types of variable. We investigated a class of aggregation operators known as copulas and their role in Sklar's theorem, which provides for the use of copulas in the formulation of joint cumulative distribution functions from the marginal CDFs of classic random variables. We then look to extend these ideas to the case of joining the marginal CDFs associated with Dempster-Shafer belief structures.

We provided a deep investigation of the use of monotonic set measures in the representation of uncertain information about a variable and its role in hard/soft information fusion. The concept of the dual of a measure was used to define the measures of opportunity and assurance associated with occurrence of an event. We observed that these generalize the ideas of possibility and necessity in possibility theory as well as the ideas of plausibility and belief in Dempster-Shafer theory. Using the measures of opportunity and assurance we are able to introduce a concept of entailment, between two uncertainty representations. We showed using the idea of entailment, how to relate possibilistic and probabilistic uncertainty and its potential application in hard/soft information fusion.

Two important aspects of the multi-source fusion problem are the representation of information provided by the sources and the formulation of the instructions on how to fuse the information provided, which we refer to as the fusion imperative. We investigated the use of a monotonic set measure as a means of representing the fusion imperative.

Aggregation plays a central role in many aspects of information fusion. We focused on a type of aggregation imperative called prioritized aggregation. We investigated two approaches to the formulation of this type of aggregation process. One of these used the prioritized aggregation operator and the second is based on an integral type aggregation using a monotonic set measure to convey the prioritized imperative. We looked at the possible role of this type of aggregation in addressing conflicting information.

Classic information fusion is generally concerned with the problem of fusing multiple pieces of uncertain information of about the same attribute to obtain a better estimate of the variable. In the case of probabilistic information the concept of correlation plays a fundamental role in allowing information about two different variables be used to improve our estimate of each of the variables. We investigate the use of a concept that we refer to as relatedness to extend this capability to the kind soft linguistic information provided by human sources.

We deeply work on the development of the use of monotonic set measures in the representation and fusion of uncertain information about the value of a variable. Under this representation we suggested that the measure of a set indicates our anticipation that the value of the variable lies in that set. We showed that for many cases of uncertainty representations, probabilistic uncertainty being a notable exception, a strong anticipation that the value of a variable lies in a given set does not preclude the possibility of a strong anticipation that it lies in the negation of the set. More generally in these cases knowledge about the anticipation of an event occurring does not imply anything about the occurrence of the negation of the event. In decision-making situations where we are interested in the occurrence of some event, we often need also to know that the negation of an event is also not anticipated. In order to provide this

type of information the concept of the dual of a measure was introduced. Using the measure of a set and its dual measure we defined the concepts of opportunity and assurance associated with occurrence of an event. We showed that these generalize the ideas of possibility and necessity in possibility theory as well as the ideas of plausibility and belief in Dempster-Shafer theory. Using the measures of opportunity and assurance we were able to introduce the concept of entailment between two uncertainty representations that generalizes Zadeh's idea of entailment between possibility distributions.

Many modern technological applications involve the aggregation of a collection of criteria to help in some selection process. Examples of this occur in information retrieval, multi–criteria decision-making, database retrieval and pattern recognition. Often these applications involve the selection from a collection of alternatives based their satisfaction to the collection of criteria. In this process we typically find the satisfaction of an alternative to the individual criteria and then aggregate these individual criteria satisfactions to obtain a fused overall score for the alternative. These overall scores are then used to select between the alternatives. The process for combining the individual criteria satisfactions is determined by what is called the aggregation imperative. The aggregation imperative is a description as to how the individual criteria satisfactions should be combined to obtain the overall score, this is typically provided by the responsible decision maker. This description can be provided in many different ways depending upon the disposition and capabilities of the relevant decision maker. Among the most common ways is a combination of mathematically and linguistically expressed instructions. An important objective of modern computational intelligence disciplines is to provide technologies to help convert these instructions into formal computational procedures. We looked at one class of aggregation imperatives, called prioritized aggregation, and obtained some formal mechanisms for implementing this type of imperative. Intuitively, the meaning of the term prioritization as used here reflects the situation where lack of satisfaction to higher priority criteria cannot be compensated for by satisfaction to lower priority criteria. In many decision processes, security has high priority in the sense that we are reluctant to tradeoff a decrease in security for a benefit in some other criteria. A linguistic formulation that often conveys this prioritized aggregation imperative is the expression "I want criterion A and if possible I also want criterion B". We investigated two formulations for implementing this type aggregation. The first was the prioritized aggregation operator. The second was an approach based on an integral type aggregation such as the Choquet integral which uses a monotonic set measure to convey the prioritized imperative.

Many environments and situations consist more and more of a combination of humans and non-human autonomous agents. We looked at the use of soft computing approaches to support the development of tools and formal mathematical concepts to enable the communication and coordination between these various heterogeneous components. An important issue that arises with such heterogeneous entities is the need to provide a common understanding of shared information, situation assessments and the goals and tasks of the specific environment. The problem is particularly acute between the human and the non-human components as they essentially employ differing communication modalities. We showed how some approaches utilizing fuzzy sets and the related theory of approximate reasoning can play an important role in helping solve this problem by providing a bridge between the types of linguistic expression and cognition that human beings use with the types of formal mathematical representations needed for the digitally based autonomous agents.

Similarity matching plays an important role in many digital systems that are used to implement fusion and other human like intelligent activities, it is fundamental to human behavior modeling. Human beings have very sophisticated ways of determining the similarity between objects. Toward the development of technologies for human behavior modeling we introduced a non-standard approach for the determination of similarity between objects, which we referred to as Hyper-matching. This approach can enable rapid similarity matching by focusing on the role of extreme or notable values in object matching. This can be particularly useful in combat environments where it is beneficial to be able to rapidly match a current situation with some similar situation for which we have some experience. In the proposed approach importance weights are associated with the relevant features used for the matching. We focus on extreme and notable attribute values by introducing an amplification of attribute importance, the more notable the feature, the more the amplification. This amplification gives these features a more important role in the matching. In this approach the weights associated with an attribute in the similarity matching become dependent on attribute value and as such this provides a type of non-linearity that allows us to focus on the fundamental notable features of the objects be compared.

Fuzzy system modeling provides a computationally intelligent method for building models of complex systems, human reasoning behavior and mathematical functions. They can be used to model intelligent fusion rules. A fuzzy systems model uses a fuzzy rule base in which the antecedent conditions of the rules are expressed using fuzzy sets. Central to the use of these models is the determination of the satisfaction, firing level, of the antecedent conditions based on information about the associated variables, the input to the fuzzy model. For the most part this input information has been expressed also using fuzzy sets. We worked on extending the capabilities of the fuzzy systems modeling technology by allowing a wider class of uncertain input information. In particular we considered the case where the input information about an antecedent variable is expressed using a measure representation of uncertain information. In providing this extension a particularly important issue is the determination of the satisfaction of a fuzzy set antecedent when the input information about the associated variable is expressed in terms of a measure. We focused on addressing this problem. After looking at some approaches for determining this firing level we provided the requirements needed by any formulation for this operation when our input information is a fuzzy set. We next introduced the idea of a measure and showed how it can be used to more generally express our knowledge about an uncertain value associated with a variable. We then generalized the requirements for any formulation that can be used to determine the satisfaction of the antecedent fuzzy set when the input information about the variable is expressed using a measure. We further provide some examples of formulations. Since a probability distribution is a special case of a measure we are able to determine the firing level of fuzzy rules with probabilistic inputs.

Decision making in situations in which there is a probabilistic uncertainty associated with the payoff that result's from the selection of an alternative is a very common task. Here each alternative is characterized by an uncertain payoff profile, a probability distribution over possible payoffs. A crucial problem here is the selection of a preferred alternative from a set of possible alternatives. While the objective is clear, select the alternative that gives the biggest payoff, the comparison of these uncertainty profiles with regard to this objective is difficult. One well-regarded method for comparing two uncertainty profiles is via the idea of stochastic dominance. While providing an intuitively reasonable paradigm for deciding which of two alternatives is preferred, stochastic dominance is a strong condition and generally a stochastic dominance

relationship between two alternatives does not exist, neither one stochastically dominates the other. In order to provide operational decision tools we looked for surrogates for stochastic dominance. These surrogates associate with each alternative a numeric value, the larger the value the more preferred, and hence always allows comparison between alternatives. An important feature of these surrogates is their consistency with stochastic dominance in the sense that if A stochastically dominates B then the surrogate value of A is larger then the surrogate value of B. We investigated a class of surrogates that we refer to as Probability Weighted Means (PWM). We determined the required properties of these PWMs. We look at a number of different examples of Probability Weighted Means.

The Dempster-Shafer belief structure provides a very useful technology for modeling an uncertain variable for which our knowledge of the probabilities is interval-valued. This is commonly referred to as second order uncertainty. Within the framework of Dempster-Shafer theory an issue that has been investigated in considerable detail is the fusion of multiple belief functions providing information about the same variable. Our interest was on a slightly different problem. Assume U and V are two random variables and our knowledge about each is expressed using Dempster-Shafer belief structures. Here our concern is not with the fusion of our information but the determination of the joint relationship between the two variables, that is we wanted to obtain the joint Dempster-Shafer belief structure in the same spirit as joint probability distributions. Toward the accomplishment of this goal we looked at the formation of the cumulative distribution function (CDF) for belief structures. We showed that these CDFs are also interval-valued and are expressible in terms of plausibility and belief measures. We then extended Sklar's theorem, which provides for the use of copulas in the formulation of joint cumulative distribution functions from the marginal CDFs for classic random variables, to the case of formulating joint cumulative distribution functions for Dempster–Shafer belief structures using copulas.

We worked at the formulation of a generalized framework for mean aggregation operators with a view toward the modeling of human cognitive aspects of information fusion. We provided an overview of mean/averaging operators. We looked at the issue of importance weighted mean aggregation. We provide a generalized formulation using a fuzzy measure to convey information about the importance of the different arguments in an aggregation. We looked at some different measures and the associated importance information they manifest. We further generalized our formulation by allowing for the inclusion of an attitudinal aggregation function. This allows us to implement many different types of aggregation including Max, Min and Median. We looked at the related problem of group consensus formation, which involves a process of getting a group of agents to agree upon a solution to some problem. Often in these types of problems each of the individual agents have their own suggested solution. We observed this problem can be seen as a type of information fusion. One important task in the formulation of a consensus is the construction of a solution as a proposed answer. This task often involves an aggregation of the different proposals provided by the individual participants. An important property of any process used in this type of aggregation is idempotency; if all the participants suggest the same solution this should clearly be a good consensus solution. Aggregation operators that manifest this idempotency are referred to as mean operators. This allowed us to use our generalized framework for implementing the mean operator and look at various formulations that are special cases of this framework. One goal here is to eventually enable the mathematical modeling of linguistically expressed mean type aggregation operations.

The paradigm of computing with words and the related theory of imate reasoning provides a framework which can be used by robot soldiers and other autonomous agents to interact with humans and reason with imprecise linguistic information. In this reasoning system information provides a restriction on the values variables can assume. One task that arises in using this approach is the formulation of joint restrictions on multiple variables from individual information about each of the variables. For example assume we are interested in an approaching enemy group. Assume one source of information tells us something about the number of men in the group and another source informs us something about how fast they are moving. In this case the joint variable would be the number of men and speed of the group and its domain values would be pairs of values, one corresponding to number and the other to speed. The fact that there is some relationship between size of a group and the speed with which it moves implies that in forming the joint variable we can get some reduction in the uncertainties of the original information. We extended the capability of the framework of computing with words to the task of forming joint variables with the introduction of the idea of perceived relatedness between variables, a concept closely related to the idea of correlation. We were particularly interested in role that information about perceived relatedness between variables can play in further restricting the possible values of joint then that simple provided by the individual constraints. We looked at the problem of joining various types of uncertain variables.